VOLUME 26 • NUMBER 3

October 1955

the RESEARCH quarterly

STEVENS RICE
UNIVERSITY MICROFILMS
313 N FIRST ST
ANN ARBOR MICH
HPR DEC P ADV

Published by THE AMERICAN ASSOCIATION FOR HEALTH, PHYSICAL EDUCATION, AND RECREATION
1201 SIXTEENTH ST., N. W., WASHINGTON 6, D. C.

RECENT PUBLICATIONS

Physical Education for College Men and Women

Complete report of the Washington Conference, Oct. 1954, sponsored by AAHPER, CPEA, and NAPECW.

48 pp. $1.00

Health Education Conference Report

Report of the National Conference on the Undergraduate Health Education Minor Program and Desirable Health Education Emphases for the Physical Education Major Program, Washington, Jan. 1955.

36 pp. $1.00

Order from

AAHPER Publications-Sales
1201 - 16th St., N.W.,
Washington 6, D. C.

Professional Contributions No. 3

Papers and reports of two annual meetings, American Academy of Physical Education, Chicago, 1953; New York City, 1954.

Includes four papers on the effects of competition in inter-school athletics.

146 pp. $1.50

RESEARCH QUARTERLY Binder

- Green leatherette cover with gold lettering
- Label-holder on backbone for volume identification
- Holds one year's issues and index or supplement
- Strong metal attachments

ONLY $2.50 each

POSTPAID

Payment must accompany order

the RESEARCH Quarterly

of the American Association for Health, Physical Education, and Recreation, Department of the National Education Association, 1201 Sixteenth St., N.W., Washington 6, D. C.

Vol. 26, No. 3 October 1955

CONTENTS

Strength Decrements from Carrying Various Army Packs on Military Marches H. HARRISON CLARKE, CLAYTON T. SHAY, and DONALD K. MATHEWS 253

Strength Decrements from Wearing Various Army Boots and Shoes on Military Marches H. HARRISON CLARKE, CLAYTON T. SHAY, and DONALD K. MATHEWS 266

Effect of Trained and Untrained Testers Upon the Admininstration of Grip Strength Tests GEORGE F. COUSINS 273

A Factor Analysis of Selected Wartime Fitness Tests GEORGE F. COUSINS 277

Steady-Pace vs. Variable Speed in High-School 220-Yard Run FRED KRONSBEIN 289

Influences of Physical Education Activities on Motor Ability and Physical Fitness of Male Freshmen CARL W. LANDISS 295

Vitamin B_{12}: A Review HENRY J. MONTOYE 308

Analysis of Results from the Kraus-Weber Test of Minimum Muscular Fitness in Children MARJORIE PHILLIPS, CAROLYN BOOKWALTER, CHARLOTTE DENMAN, JANET McAULEY, HILDA SHERWIN, DEAN SUMMERS, and HELEN YEAKEL 314

Measurement of Kinesthesis M. GLADYS SCOTT 324

Emotional Responses of Boys to Little League and Middle League Competitive Baseball ELVERA SKUBIC 342

Game Results as a Criterion for Validating Basketball Skill Test FRANCIS STROUP 353

Emotion in Fitness Tests PHILIP V. WELLS 358

NOTES AND COMMENTS—**Professional Courses in the Prevention and Treatment of Athletic Injuries at American Universities and Colleges,** Lane B. Blank, 363; **Predictive Ability of a Selected Basketball Test,** Clifford A. Boyd, James R. McCachren, and I. F. Waglow, 364; **Comments on the article by Alfred W. Hubbard and Charles N. Seng, "Visual Movements of Batters," Research Quarterly, 25:42-57 (March 1954),** Arthur T. Slater-Hammel, 365; **Rebuttal to above comments on "Visual Movements of Batters,"** Alfred W. Hubbard, 366.

Research Abstracts 369

Guide to Authors 375

Copyright 1955, by the American Association for Health, Physical Education and Recreation National Education Association

Carl A. Troester, Jr., Executive Secretary

Assistant Secretaries

Elizabeth Avery, Consultant in Health Education

Rachel E. Bryant, Consultant in Physical Education and Girls and Women's Sports

Jackson M. Anderson, Consultant in Recreation and Outdoor Education

Eugene H. Sloane, Director of Publications and Public Relations

●

Ella H. Wright, Editor

Kiyoko A. Wittenburg, Assistant Editor

Jean W. Sloan, Circulation Manager

George F. Anderson, Associate Executive Secretary in charge of Advertising

Ames A. Castle, Sports and Industrial Relations Executive

●

Board of Associate Editors

Karl Bookwalter
Marion R. Broer
H. Harrison Clarke
Thomas K. Cureton, Jr.
Anna Espenschade
Franklin M. Henry
Pauline Hodgson
James H. Humphrey
Paul Hunsicker
Leslie W. Irwin
Warren R. Johnson
H. F. Kilander
Joy W. Kistler
Hyman Krakower
Lynn W. McCraw
Kenneth D. Miller
Henry J. Montoye
Laurence E. Morehouse
Marjorie Phillips
Laurence Rarick
M. Gladys Scott
John H. Shaw
Frank D. Sills
Wesley M. Staton
D. B. Van Dalen
C. Etta Walters
Allen E. Weatherford

Published in March, May, October, and December by the American Association for Health, Physical Education, and Recreation, 1201 Sixteenth Street, N.W., Washington 6, D. C. Subscription available to libraries, institutions, and departments at $5.00 per year. Of the Professional Membership ($10.00 per year), $3.00 covers subscription to the *Quarterly*. Of the Student Professional Membership ($5.00 per year), $1.50 is for subscription to the *Quarterly*. *Subscriptions to individuals through membership only.* Single copies $1.25.

Send membership dues or subscriptions to Circulation Dept., 1201 Sixteenth Street, N.W., Washington 6, D. C. Editorial Office, 1201 Sixteenth Street, N.W., Washington 6, D. C. Entered as second-class matter at the Post Office at Washington, D. C., under the act of March 3, 1879. Additional entry at Baltimore, Md.

Manuscripts are to be sent to the Editorial Office and are expected to follow the form prescribed in the Guide to Authors that appears each year in the October issue of the Research Quarterly.

Strength Decrements from Carrying Various Army Packs on Military Marches[1]

H. HARRISON CLARKE
University of Oregon, Eugene, Oregon

CLAYTON T. SHAY AND DONALD K. MATHEWS
Springfield College, Springfield, Massachusetts

Abstract

Two independent pack-carrying studies were conducted in which the strength losses of involved muscles from nine 7.5-mile marches were determined. In the first series of marches, the rucksack had a clear advantage: not until the weight was 61 pounds were strength losses appreciable. Contrarily, in the second series, marching with the rucksack resulted in greater strength losses than with the combat pack. The difference is attributed to the superior physical condition of the subjects on the first rucksack march. The muscle groups showing greatest strength losses for the nine marches were: trunk extensors in eight marches; hip extensors in seven marches; and knee flexors in six marches. The neck extensors appeared in two of the three marches of the second series, but were not tested in series one.

HIKING AND CAMPING are of increasing importance to physical education and recreation. If advanced instruction is to be given in these subjects, it should be based upon research information. This study deals with certain specific problems of importance in hiking with a pack, namely the determination of what muscle groups are most subject to fatigue, the degree of local fatigue, and the relative influence of pack position and load.

These problems are also of interest to the military services. Throughout history, armies have been faced with the problems of supplying soldiers with clothing and personal equipment, food and shelter, weapons and ammunition. The problem is relatively simple when soldiers are in encampment; on the march and in combat, it is much more difficult. The expanding scale of military operations and changes in technology have increased greatly the complexity of this situation. In the solution of supply, necessarily the infantryman has been forced to transport essential and immediately needed items of equipment on his person.

The foot soldier of early armies, such as the Roman legionnaire as well as the infantryman of France, Russia, Germany, and Japan prior to 1900, carried loads of 60 to 65 pounds and more. Frequently, these soldiers reached the battlefield in such a fatigued state that effective combat was impossible

[1]This study was conducted at Springfield College under subsidy from the Climatic Research Laboratory, Quartermaster Corps, Department of the Army, Natick, Massachusetts. Acknowledgments are expressed to the following graduate assistants who assisted in the collection and analyis of the necessary test data: Gavin H. Carter, David H. Clarke, Karl K. Klein, William Mahon, Barry Thompson, and James E. Torpey.

(2). As has been true of modern wars, much of the equipment thus carried was thrown away along the route of march. The soldier in combat frequently discarded everything he could do without with no particular concern as to what happened to it.

Much has been written and said on the problem of what, where, and how much the soldier should carry. Despite all the controversy over the centuries, little scientific evidence is available for the solution of this problem, decisions, for the most part, being based upon opinion. The lack of factual data has been the single greatest deterrrent to the development of a completely functional system of load carrying. During recent years, the Quartermaster Corps, United States Army, has conducted investigations into the design of special packs and into ways of lightening the weight of essential items of equipment. Through its Climatic Research Laboratory, the effects of pack carrying on the human organism are being investigated under experimental conditions.

A review of the physiological research on pack carrying has been presented by Bailey and McDermott (1). In the main, these studies have two characteristics: (a) The evaluative processes have been based largely on metabolic determiners; and to a less extent, on electromyographic techniques, center of gravity measurements, and the judgment of subjects. (b) The experimental procedures have been confined mostly to treadmill marching, thus without reference to the field conditions encountered by soldiers or by hikers over uneven terrain. These reviewers conclude that the results so far obtained are meager and frequently inconclusive.

Shortcomings of the investigations reported are as follows: (a) Many studies were based upon limited observations of a small number of subjects. (b) The state of training and other descriptive measures of the subjects were not known. (c) Psychological motivational factors in pack-carrving performance were ignored. (d) Energy cost studies do not adequately reflect the fatigue of local muscle groups. (e) Only muscles directly underlying the pick-up electrodes can be investigated by electromyography; and there is no differentiation between action and stabilizing muscles. (f) There was no reliable measure used for determining degrees of muscular contraction. (g) The center of gravity studies were conducted largely from a standing rather than a marching position.

Background of the Study

The present investigation into the effect of marching with military packs is based upon two major factors, as follows:

1. The hypothesis for the evaluative process is that muscles under physical stress may weaken when carrying loads. The validity of this hypothesis was clearly demonstrated in another study reported by the authors (4): The subjects fatigued their elbow flexor muscles by ergographic exercise with a load equal to three-eighths of the strength of these muscles; mean loss in strength 30 sec. after such exercise was approximately 32 per cent and strength recovery was not complete three-fourths of an hour later.

In the present study, therefore, the strength loss from marching with various loads carried in different positions on the body and in different army packs was determined for the muscles primarily involved. The strength decrement technique identifies muscle groups affected by localized fatigue; and also may indicate general body fatigue by combining the strength losses of all muscles so affected.

2. The study was designed for field conditions with the actual load and load distribution on the body of the rifleman. Thus, it is recognized that the foot soldier in the field is not limited to marching on a smooth surface (as on a treadmill), but must often travel long distances over rough terrain; and does not carry the entire weight in a lump on the back (as frequently done in such studies), but distributes parts of the load to other areas of the body. These same observations are true for hikers on outdoor trails.

Research Procedures

FIELD CONDITIONS

1. The field course was 1.25 miles long over moderately rough and hilly terrain. Each march was 7.5 miles long, requiring six repititions of the course.

2. The rate of march was 2.5 miles in 50 min.; a ten-minute stop permitted time for re-testing the subjects during the march. A pace-control system was devised to insure adherence to this rate of march.

3. The subjects were provided with regulation army clothing, combat boots, equipment, packs, and M-1 rifles. The subjects were dressed and the load distributed on the person, with the exception of the packs which were varied, in accordance with Army regulations. This arrangement is shown in Figure I, with combat pack carried low on the back.

4. An exploratory march and six experimental marches were conducted. The specific conditions for the six experimental marches were as follows:[2]

March No. 1: Combat pack without rifle; weight 29 lb.; pack carried high on back; winter clothing.

March No. 2: Same as March No. 1, except with rifle; weight, 41 lb.

March No. 3: Same as March No. 2, except with pack carried low on back.

March No. 4: No pack, load, or equipment; summer fatigues.

March No. 5: Rucksack with rifles and grenades; weight, 41 lb.; summer fatigues.

March No. 6: Same as March No. 5, except with 61 lb. (which equals the field-load weight of the rifleman, but not the pack carried).

STRENGTH TESTING

1. *Strength Tests:* The muscle groups tested[3] were those considered to be: (a) under direct stress in supporting the packs, (b) involved in stabilizing the rest of the body in relation to the pack, (c) primarily activated in marching,

[2]The weights given include all equipment carried by the subjects, not alone the weight of the pack.

[3]For assistance in the selection of these muscle groups, acknowledgments are made to Dr. Earl C. Elkins, Section on Physical Medicine, Mayo Clinic, and to Dr. F. A. Hellebrandt, Head, Department of Physical Medicine and Rehabilitation, University of Illinois Research and Educational Hospitals.

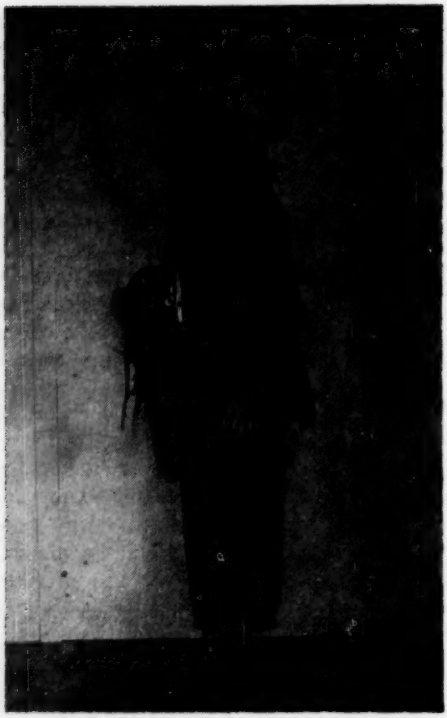


FIG. I. Military Equipment: Pack Carried Low on Back.

and (d) affected by interference with circulation caused by pressure from pack straps. The following 11 muscle groups were tested for strength at the beginning and at the end of each 2.5 miles of march, in the order given: right grip, left grip, ankle plantar flexors, pectorals, knee extensors, knee flexors, hip flexors, hip extensors, trunk flexors, trunk extensors, and shoulder elevators. With the exception of grip strength, which was measured with a manuometer, the strength tests were of the cable-tension type, previously developed in the Springfield Laboratory.[4]

2. *Strength Decrement Index:* The Strength Decrement Index (SDI) was the basic score used in this research. It is computed as follows:

$$SDI = \frac{S_b - S_a}{S_b} \times 100$$

In which: S_b = Before march strength
S_a = After march strength

[4]Detailed data and graphs on this research will be found in references 5 and 6. Also, included in these references are descriptions of all tests and special devices used in this study.

3. *Testers:* Prior to the actual collection of research data, the graduate students serving as research assistants were carefully trained in the strength testing techniques. They were required to demonstrate testing competence by obtaining objectivity coefficients by the test-retest method with different testers approximately equal to those established for the tests. The coefficients thus obtained ranged from .86 to .98.

SUBJECTS

The subjects were 30 Springfield College students majoring in physical education. The age range was between 18 and 26 years. The following median scores will further describe this group: height, 69.9 in.; weight, 167 lbs.; Physical Fitness Index, 104; and Strength Index, 3013.

Before the pack-carrying experiment commenced, the subjects were taken to the marching area in small groups and were oriented to the purpose of the study. Then, they were taken over the field course with packs, so as to become familiar with the trail and the rate of march which would be necessary. The strength testing sequence was explained and they went through the process of test administration.

Results

A summary of the results of the pack-carrying study appears in Table 1. Both mean SDI's and t ratios above 1.50 (representing the significance between pre- and post-march means) are given for each muscle group on each

TABLE 1

Summarization of Strength Decrement Indices for the Six Experimental Marches

Marches	Mean	Muscle Groups							
		Shoulder Elevators	Trunk Flexors	Trunk Extensors	Hip Flexors	Hip Extensors	Knee Flexors	Knee Extensors	Ankle Plantar Flexors
1. 29-lb. high combat pack	SDI		9.31	11.19	3.64	6.34	13.78	5.19	3.57
	t		1.72	2.00	2.29	1.74	4.20	1.53	1.65
2. 41-lb. high combat pack	SDI		3.18	3.70	4.21	5.04	13.01	5.48	3.64
	t		1.71	2.50	2.50	2.30	4.59	3.54	2.18
3. 41-lb. low combat pack	SDI	10.28	4.30	8.09	5.97	4.28	10.30	4.99	4.49
	t	5.30	1.57	1.75	2.16	2.43	3.13	2.00	2.19
4. No load	SDI		7.84	6.40					4.68
	t		2.00	1.71					2.13
5. 41-lb. rucksack	SDI				4.36	3.46			3.29
	t				1.60	1.62			2.14
6. 61-lb. rucksack	SDI	10.10	8.08	14.64	5.35	5.75	6.46		6.58
	t	2.79	4.67	4.53	2.40	1.67	2.47		2.86

march.[5] The most significant figures on this table are the *t's*. The mean SDI's are affected by any unusual situation, causing an extreme score, which could inadvertently happen during a given march. Such extreme scores increase means disproportionately to the importance of single cases. The t is not so affected, as the extreme scores also produce greater standard deviations, which in turn are reflected in larger standard errors for these measures.

Eight muscle groups, all except the right and left grips and the pectorals, had t's above 1.50 for at least three of the marches. The exceptions noted did not produce statistically significant results for any of the pack carries, so are eliminated from Table 1 and from further consideration in this report. The grip strength tests were included in this research as it was thought that the pressure from the pack straps in carrying military loads might reduce their strength as a consequence of interference with blood circulation to the arms and hands. However, such results were not obtained from these tests.

Observations of the strength losses resulting from the various marches are as follows:

1. Carrying a 41-lb. load with the rucksack resulted in the least over-all strength loss of the muscle groups tested. Only three muscle groups, all located in the lower body (marching muscles), had strength losses which were statistically significant near and beyond the .05 level of confidence: hip flexors, hip extensors, and ankle plantar flexors.

2. With the rucksack, also, the effect on total strength loss with the 61-lb. load was approximately the same as for the combat pack with 41 lbs. carried low on the back. However, the proportionate amount of loss for the different muscle groups varied: with the rucksack, greater loss was found for the trunk flexors and extensors; with the combat pack, the larger decrements were for the shoulder elevators and the knee extensors.

3. Very little difference in over-all strength loss was found between the 41-lb. high and low combat pack marches, although there were differences in the muscles affected. The greater losses were: knee flexors and extensors for the high pack; and shoulder elevators for the low pack.

4. With no load, three muscle groups were involved, the trunk flexors and extensors and the ankle plantar flexors. From observing the results of the other marches, loads up to 41 lb. did not affect the mean SDI's of these muscle groups beyond those involved in marching the 7.5 miles. However, when the load was increased beyond this weight, definite additional stress was found, as shown for the 61-lb. rucksack march.

---

[5]Significant t's, when probability is confined to one tail of the normal distribution, are: 1.70 at the .05 level of confidence; and 2.47 at the .01 level. The standard error formula for correlated groups was used in these computations. Inasmuch as the subjects utilized in this study did not constitute a random sample, but, rather, were physically superior to the normal population, the t-ratios should be interpreted cautiously.

5. The knee flexor muscles had the greatest total strength loss on the six experimental marches. However, the strain on these muscles was centered primarily in the first three and last marches, with t's of 4.20, 4.59, 3.13, and 2.47 respectively. As was subsequently found, the amount of knee flexion loss for the final march, when the subjects were in their best marching condition, is the most significant single SDI in correlating with motor-physical fitness tests. The correlations were: —.53, with Strength Index; —.49, with AAF Physical Fitness Test; —.45, with Navy Standard Physical Fitness Index. In interpreting these correlations, r's of .35 and .45 are significant at the .05 and .01 levels of confidence respectively.

6. For the shoulder elevator muscles, only two marches produced statistically significant mean SDI's, the 41-lb. low combat pack and the 61-lb. rucksack. In both of these marches, the load is carried low but with straps over the shoulders, thus resulting in a long pull on the shoulder muscles. The effect of this pull was not noted with the rucksack (41 lb.) until the weight became much greater (61 lb.), which may be accounted for by the brace construction of this pack providing support at the hips.

7. The knee extensors had the least mean strength losses of the eight muscle groups. They followed much the same pattern, but in lesser degree, as did the knee flexors, except no significant mean strength loss was found for the 61-lb. rucksack march. This exception may be due to the better condition of the subjects on the final march in the series.

8. Mean strength losses of over 10 per cent were found in seven muscle groups during the marches. The SDI's were as follows: 14.64, trunk extensors with 61-lb. rucksack; 13.78 and 13.01, knee flexors with 29-lb. and 41-lb. high combat packs respectively; 11.19, trunk extensors, with 29-lb. high combat pack; and 10.28 and 10.10, shoulder elevators, with 41-lb. low combat pack and 61-lb. rucksack respectively.

Tables, graphs, and discussions of the mean SDI's for each muscle group at 2.5, 5.0, and 7.5 miles of marching are contained in the original reports by the authors (5, 6). The presentation of so much detailed information is not made in this article. To show the nature of these data, however, the mean strength decrements of the knee flexor muscles are presented in Figure II.

Nearly the full strength decrement effect on the knee flexor muscles was felt in the first 2.5 miles of the first two marches, the 29-lb. and 41-lb. high combat packs; they continued at about the same level for the balance of the marches. With the 41-lb. low combat pack, the decrement line rose sharply at 2.5 miles (about two-thirds of the total loss), after which it continued as a straight line rise but at a lessened rate. The 61-lb. rucksack march produced an increase in strength at 2.5 miles, followed by a rapid decline to 7.5 miles; however, the final SDI of 6.46 is not significant at the .05 level of confidence. The decrements for the other two marches—no pack and 41-lb. rucksack—were slight and not statistically significant.

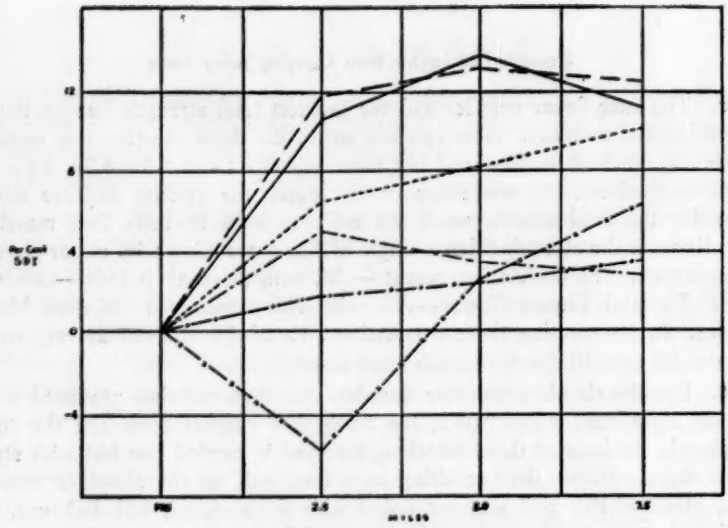


FIG. II. Knee Flexion Mean SDI's Pack-Carrying (Data: 1952-53).
Pack: ——— = 29 lb. high; — — — = 41 lb. high; · · · · = 41 lb. low;
— · — · — = no pack; — · — · — = 41 lb. rucksack; — x — x — = 61 lb. rucksack

Second Pack-Carrying Study[6]

In the preceding pack-carrying study, no effort was made to control the physical conditioning effect on the subjects from repeated 7.5 mile marches. That such conditioning took place is readily concluded from the following two sources of evidence:

1. *Increase in strength of the muscle groups tested:* The mean pre-march strength scores for each muscle group tested were computed before the first and last marches. Thus, five 7.5-mile marches intervened between tests. The results appear in Table 2.

TABLE 2

Increase in Strength of Involved Muscles as a Result of Five 7.5-Mile Military Marches

Muscle Group	Means before		Diff.	σD	t
	First March	Last March			
Trunk extensors	124.6	164.5	39.9	8.55	4.67
Trunk flexors	121.6	157.9	36.3	8.82	4.12
Shoulder elevators	452.3	524.0	71.7	22.47	3.19
Hip flexors	155.5	170.1	14.6	5.99	2.44
Ankle plantar flexors	370.8	388.9	18.1	11.98	1.51
Hip extensors	111.4	113.2	1.8	6.26	.29
Knee flexors	118.2	119.5	.7		
Knee extensors	217.0	217.8	.8		

[6]For this phase of the study, acknowledgments are made to the following graduate students who assisted in the collection and analysis of the necessary test data: Douglas A. Bailey, Theodore Fabyan, Frank A. Godin, Allan H. Goodyear, Robert E. Hodgdon, Karl K. Klein, Adrian J. Marinelli, Larry E. McClements, Frank M. Robinson, David M. Seamans, and Vincent Trotta, Jr.

The increase in strength of the involved muscles was definite. The differences between the means for three muscle groups were statistically significant beyond the .01 level of confidence: trunk extensors (39.9 lb.), trunk flexors (36.3 lb.), and shoulder elevators (71.7 lb.). The gain in mean strength for the hip flexors (14.6 lb.) was nearly significant at the .01 level; and the mean gain for the ankle plantar flexors (18.1 lb.) was noticeable, but did not reach the .05 level.

Furthermore, the over-all percentage in mean loss of strength was less and generally occurred later on the final march (7.5 miles) than on the first march (2.5 to 5.0 miles). This observation, however, should be interpreted cautiously as the conditions were not identical, the packs being different for the two marches. However, the load was 20 lb. heavier for the final march, which might logically have caused greater strength losses.

2. *Increase in motor-physical fitness test scores:* Before the exploratory march and after the last experimental march (seven marches thus intervening), several motor-physical fitness test batteries were administered to the subjects. The mean scores for this testing appear in Table 3.

TABLE 3

Comparison of Means on Motor-Physical Fitness Tests Before and After Seven Military Pack-Carrying Marches

Motor-Physical Fitness Tests	Initial Mean	Final Mean	Diff.	σD	t
1. Army Physical Efficiency Test	253	281	28	4.6	6.1
2. AAF Physical Fitness Test	56	59	3	.6	5.0
3. Navy Standard Physical Fitness Test	260	275	15	3.0	5.0
4. Physical Fitness Index	104	113	9	2.0	4.5
5. Strength Index	3084	3407	323	142.0	2.3
6. McCloy's Arm Strength Score	581	591	10	6.1	1.6

The means for all motor fitness tests increased during the pack-carrying study. All increases but two were significant beyond the .01 level of confidence. The exceptions are: Strength Index with a t of 2.3, which is nearly significant at the .01 level; and McCloy's Arm Strength Score with a t of 1.64, which nearly reaches the .05 level. All three of the armed forces motor fitness tests showed definite gains, especally the Army Physical Efficiency Test, which had a t of 6.1. The final mean of 113 on the Physical Fitness Index test is a high score, as 115 is the third quartile for normal populations.

PROCEDURES

In the above study, the physical condition of the subjects might logically be a factor in the strength decrement results from the pack-carrying marches. Inasmuch as each subject executed a set performance (marching 7.5 miles at a rate of 2.5 miles in 50 minutes), improved condition could be reflected in less strength loss. Had the performance been competitive (i.e., to march the course in the shortest possible time), the SDI's could have remained much

the same, with improved condition being expressed in faster time. In fact, this latter hypothesis has been verified in a recent swimming experiment conducted at the University of Oregon (7).

Therefore, because physical conditioning of the subjects resulting from repeated marches was uncontrolled in the initial research, a second study was undertaken which attempted to neutralize this effect on pack-carrying evaluation. Two of the former packs were included: the combat pack carried high on the back and the rucksack. In addition, a Quartermaster experimental pack, designed to locate most of the weight around the waist, was included.

All weights were 41 lb. As before, the marching distance was 7.5 miles and the rate of march was 2.5 miles in 50 minutes. The subjects were 40 Springfield College students. Statistically significant t's for this N are: 1.68 at the .05 level of confidence; and 2.42 at the .01 level. The strength tests utilized in their order of administration were: shoulder elevation, neck extension, knee flexion, knee extension, hip flexion, hip extension, trunk flexion, and trunk extension.

The conditioning effect on the strength decrement of the muscles resulting from the marches was equalized by employment of the following Latin-square distribution, with the subjects divided into three groups of 13, 13, and 14 subjects, designated A, B, and C:

March	Combat Pack	Rucksack	Experimental Pack
First	A	B	C
Second	B	C	A
Third	C	A	B

RESULTS

The strength decrements for the marches carrying the three military packs are summarized in Table 4. As in Table 1, both the SDI's and the t's for the difference between pre-march means and the means for the greatest SDI during the march are presented.

TABLE 4

Summarization of Strength Decrement Indices for the Three Pack-Carrying Marches

Muscle Groups	Combat Pack		Experimental Pack		Rucksack	
	SDI	t	SDI	t	SDI	t
Shoulder elevators	8.99	2.68	6.03	1.54	4.35	1.86
Neck extensors	11.25	4.19	3.72	.88	7.53	3.11
Knee flexors	9.38	2.62	6.60	1.13	10.59	3.46
Knee extensors	5.85	1.21	2.09	.65	7.28	2.31
Hip flexors	1.42	.60	3.60	1.18	4.35	2.12
Hip extensors	4.59	1.80	6.20	1.97	3.88	2.39
Trunk flexors	2.16	.70	4.15	1.39	7.73	3.81
Trunk extensors	6.40	2.94	8.43	2.22	9.12	2.19

Observations of the strength losses occurring on the various marches are as follows:

1. Carrying the 41-lb. load by means of the QM experimental pack resulted in the least over-all strength loss of the muscle groups tested; the total

was 10.96. Only two muscle groups were significantly affected in this march, between the .05 and the .01 levels: the trunk extensors with a t of 2.22 (8.43% loss); and hip extensors with a t of 1.97 (6.2% loss). The strength loss for the shoulder elevators was noticeable but did not reach the .05 level (t of 1.54).

2. With the combat pack carried high, five muscle groups exhibited significant losses in strength, four of which were beyond the .01 level of confidence. The muscle groups were: neck extensors with a t of 4.19 (11.25% loss); trunk extensors with a t of 2.94 (6.40% loss); shoulder elevators with a t of 2.68 (8.99% loss); knee flexors with a t of 2.62 (9.38% loss); and hip extensors with a t of 1.80 (4.59% loss). The total t was 16.74, which is 53 per cent greater than for the experimental pack.

3. In this series, the march with the rucksack resulted in the greatest percentage of loss, approximately twice as much as for the experimental pack and about 25 per cent greater than for the combat pack. All eight muscle groups tested were significantly affected by strength losses.

4. The hip flexors and the trunk extensors were the only muscle groups significantly affected on all three experimental marches. Those muscle groups with significant differences on two marches were: shoulder elevators, neck extensors, and knee flexors.

5. The neck extensor muscle group was added to the battery of strength tests in this study, due to discomfort in this area experienced by the subjects in the earlier series. Considerable stress on these muscles was found for the combat pack and the rucksack, with t's of 4.19 (11.25%) and 3.11 (7.53%) respectively. For the QM experimental pack, however, the strain on this muscle group was negligible (t of .88).

Discussion

Two independent pack-carrying studies were conducted in which the strength losses of involved muscles from 7.5-mile military marches were determined. In the first study, an exploratory and six experimental marches were held. During these marches, the subjects improved considerably in the strength of the muscles tested and in motor-physical fitness. No control of this factor was exercised, so that the conditioning benefits from the marches favored the last ones in which the rucksack was used. In the second study, the conditioning effect was equalized by employment of a Latin-square distribution. However, only three marches were held in this series, so the subjects did not achieve conditioning improvement to the same degree as in the first study. In fact, the condition of these subjects from marching was about at the level of the first experimental march (actually the second march) of the first series.

In the first series of marches, the rucksack had a clear advantage in carrying military loads: not until the weight was 61 lb. were strength losses appreciable. Contrarily, in the second series, marching with the rucksack resulted in greater strength losses than with the combat pack, both loads being

41 lb. The difference is attributed to the condition of the subjects on the two marches as discussed above. This contention is further supported by the observation that the SDI's for the rucksack march with 61 lb., when the subjects were in their best marching condition, are comparable to the SDI's for the rucksack with 41 lb., when the subjects were unconditioned for marching. For the second series of marches, the QM experimental pack was clearly superior, resulting in considerably less strength decrements when marching 7.5 miles with a 41-lb. load.

In the first series of marches, the shoulder elevator muscles did not show a significant strength loss when the combat pack was carried high on the back; in the second series, the loss of strength in these muscles was significant at the .01 level of confidence. The subjects were in about the same stage of marching condition. This discrepancy can be explained by the order in which the tests were administered after marching: in the first series, these muscles were tested last, thus permitting about eight minutes of recovery time; in the second series, they were tested first with very little rest of these muscles allowed. Other similar discrepancies in SDI differences between the two series arise from like situations.

A listing of the muscle groups most seriously affected with strength losses on the various marches in the two series (t's of 3.00 and over) is as follows:

First Series

29-lb. high combat pack: knee flexors (13.78%).
41-lb. high combat pack: knee flexors (13.01%) and knee extensors (5.48%).
41-lb. low combat pack: shoulder elevators (10.28%) and knee flexors (10.30%).
No load: none.
41-lb. rucksack: none.
61-lb. rucksack: trunk flexors (8.08%) and trunk extensors (14.64%).

Second Series

41-lb. high combat pack; neck extensors (11.25%).
41-lb. rucksack: trunk flexors (7.73%), knee flexors (10.59%), and neck extensors (7.53%).
41-lb. QM experimental pack: none

The muscle groups showing greatest strength losses for the nine marches in the two series were: trunk extensors in eight marches; hip extensors in seven marches; and knee flexors in six marches. The ankle plantar flexors were included in the first series only; as they showed significant strength losses on all six of these marches, it is safe to classify them among the frequently affected muscles. The neck extensors appeared in two of the three marches in the second series, but were not tested in series one. Of the other muscle groups, the trunk flexors and the hip flexors were significantly affected in five marches each; the shoulder elevators and knee extensors in four each.

Conclusions

Observations on the strength losses resulting from the various marches with military packs performed in this study are as follows:

1. Military marches with the Quartermaster experimental pack resulted in the least muscular fatigue, as reflected in Strength Decrement Indices of the

muscle groups involved. Only strength losses in two muscle groups—trunk extensors and hip extensors—were statistically significant.

2. In the second series of marches, when the subjects were unconditioned for marching, the combat pack carried high on the back was superior to the rucksack. Three muscle groups were significantly affected by strength loss with this pack: neck extensors, trunk extensors, and shoulder elevators.

3. When conditioned, the subjects experienced much less strength loss in marching with military packs, as indicated by the lower Strength Decrement Indices resulting from the seventh march of the first series. On this march, the subjects carried a rucksack with a 61-lb. load. All of the muscle groups tested, except the knee extensors, experienced significant strength losses from this march, with the trunk flexors and extensors having the greatest proportionate loss.

4. Very little difference existed in over-all strength loss resulted between carrying the combat pack high or low on the back. However, the muscles affected differed somewhat: the shoulder elevators had greater loss with the low pack; and the knee flexors and extensors, with the high pack.

REFERENCES

1. Bailey, Theodore L., and William H. McDermott. *Review of Research on Load Carrying.* Tentage and Equipment Series Report No. 9, Office of the Quartermaster General, Department of the Army, 1952.
2. Carre, Lt. *Historical Review of the Load of the Foot Soldier.* Tentage and Equipment Series Report No. 8, Office of the Quartermaster General, Department of the Army. (Translated by Captain P. L. Miles for publication in the *Infantry Journal,* 1908.)
3. Clarke, H. Harrison. A Manual: Cable-Tension Strength Tests. Brown-Murphy Co.: Chicopee, Mass., 1953.
4. ———, Clayton T. Shay, and Donald K. Mathews. Strength Decrement Effects of Exhaustive Efforts of the Elbow Flexor Muscles. *Archives of Physical Medicine and Rehabilitation,* **35**: 560 (Sept. 1954).
5. ———. *Strength Decrement Indices from Military Marches: Comparison of Army Packs and Army Boots and Shoes.* Progress Report, Climatic Research Laboratory, Army Quartermaster Corps, Natick, Mass., 1954.
6. ———. *The Effects of Pack Carrying on the Muscular Strength of Involved Muscles.* Progress Report, Climatic Research Laboratory, Army Quartermaster Corps, Natick, Mass., 1953.
7. Davis, Jack F., *Effects of Training and Conditioning for Middle Distance Swimming upon Measures of General Physical Fitness, Gross Strength, Motor Fitness, and Strength of Involved Muscles.* Unpublished doctoral dissertation, University of Oregon, 1954.

(Submitted 10/5/54)

Strength Decrements from Wearing Various Army Boots and Shoes on Military Marches[1]

H. HARRISON CLARKE
University of Oregon
Eugene, Oregon

CLAYTON T. SHAY AND DONALD K. MATHEWS
Springfield College
Springfield, Massachusetts

Abstract

The strength loss effects on the muscles of the lower leg and ankle resulting from wearing three types of army boots and shoes on 7.5 mile marches carrying combat packs were determined. Military marches with the combat boot resulted in the least fatigue of the muscle groups studied; the rubber insulated boot was next favored; and strength loss with the quarter-cut shoe was the greatest. The muscle group most affected by proportionate strength losses for the three footwear marches was the ankle everters; and, then, in order: the ankle plantar flexors, the ankle inverters, and the ankle dorsi flexors.

AMONG THE INCREASING interests of physical education and recreation today are hiking and camping in the out-of-doors. The long hike with pack over woodland trails has become an enjoyable leisure-time activity for many people. A problem of vital importance to personal comfort on such hikes is the type of footwear worn. This study deals with an evaluation of the muscular fatigue effects from wearing various boots and shoes on a standardized hike.

In another study by the authors (2), the strength decrements of muscles primarily involved in carrying army packs on military marches were determined. The marches were 7.5 miles in length; and the rate of march was 2.5 miles in 50 min. The subjects were supplied with regulation army clothing, equipment, and rifles. The packs were: 41-lb. combat pack carried high on the back, the same pack carried low on the back, 41-lb. rucksack, 61-lb. rucksack, and 41-lb. experimental pack (load located primarily around the waist). The strength decrements of 12 muscle groups were measured: right grip, left grip, shoulder elevators, pectorals, neck extensors, trunk flexors and extensors, hip flexors and extensors, knee flexors and extensors, and ankle plantar flexors. The muscle groups under greatest stress from fatigue were identified for each pack; and the over-all effect was determined by combining the strength losses for all affected muscles. The march with the experimental pack resulted

[1]This study was conducted at Springfield College under subsidy from the Climatic Research Laboratory, Quartermaster Corps, Department of the Army, Natick, Massachusetts. Acknowledgments are made to the following graduate assistants who assisted in the collection and analysis of the necessary test data: Douglas A. Bailey, Theodore Fabyan, Frank A. Godin, Allan H. Goodyear, Robert E. Hodgdon, David M. Seamans, and Vincent Trotta, Jr.

in the least fatigue, as tested by strength decrements, followed in order by the combat pack and the rucksack. The physical condition of the subjects improved considerably as a result of repeated pack-carrying marches; and this factor was reflected in lessened strength losses for the late marches in the series.

The present investigation was designed to compare the strength loss effects on the muscles of the lower leg and ankle resulting from wearing three types of army boots and shoes on a uniform pack-carrying march.

Research Procedures

The research procedures applied in this investigation are similar to those described for the evaluation of army packs. Inasmuch as this report is also published in the *Research Quarterly* (2), a summary of this process only will be given here, as follows:

1. The following types of army footwear were included in the study: combat boot, rubber insulated boot, and quarter-cut shoe. The subjects completed three marches, wearing a different boot or shoe each time. The conditioning effect on the strength of the muscles resulting from repeated marches was equalized by employment of the following Latin-square distribution, with the subjects divided into three equal groups, designated A, B, and C:

March	Combat Boot	Insulated Boot	Quarter-cut Shoe
First	A	B	C
Second	B	C	A
Third	C	A	B

2. The marching distance was 7.5 miles over moderately rough terrain. The rate of march was 2.5 miles in 50 min.; a 10-min. stop at the end of each 2.5 miles permitted time for strength re-testing periodically during the marches. The same pace-control system used in the pack-carrying study insured adherence to the rate of march.

3. The combat pack, located high on the back, was carried on all marches. The total weight of the load was 41 lb., which includes the pack, rifle, ammunition, and other equipment carried by the rifleman. The subjects were dressed in accordance with army regulations, and the load was distributed on the person as it would be carried in the field.

4. The subjects were 29 Springfield College students majoring in physical education. They ranged in age from 18 to 26 years. The following median scores will further describe this group: height, 69 in.; weight, 155 lb.; Physical Fitness Index, 107; Strength Index, 2900.

5. Prior to the actual collection of test data, the graduate assistants serving as research assistants were carefully trained in testing techniques. They were required to demonstrate testing competence by obtaining objectivity coefficients by the test-retest method with different testers approximately equal to those established for the tests.

6. The strength tests were administered each time in the same order, from weak to strong movements, as follows: ankle inverters, ankle everters, ankle dorsi flexors, and ankle plantar flexors. Clarke's cable-tension tests (1) were utilized to measure the strength of these muscle groups.

7. The Strength Decrement Index (SDI) was computed for each muscle group on each march at 2.5, 5.0, and 7.5 miles. This index is computed as follows:

$$SDI = \frac{S_b - S_a}{S_b} \times 100$$

In which: S_b = Before march strength
S_a = After march strength

8. In determining the statistical significance of the difference between pre-march and post-march means, the null hypothesis based upon "Student's" distribution was used. For 29 cases, significant t's, when probability is confined to one tail of the normal distribution, are: 1.70 at the .05 level of confidence; and 2.47 at the .01 level (3). Inasmuch as the subjects utilized in this study did not constitute a random sample, but, rather, were physically superior to the normal population, the t-ratios should be interpreted cautiously.

Results

ANKLE INVERTERS

The SDI's obtained on the marches for the ankle inverter muscles are shown in Figure I. Observations on these results are:

1. The march in which the combat boot was worn had the least effect on strength loss of this muscle group. In fact, they increased in strength scores at 2.5 and 5.0 miles; and had just reached their pre-march mean at 7.5 miles.

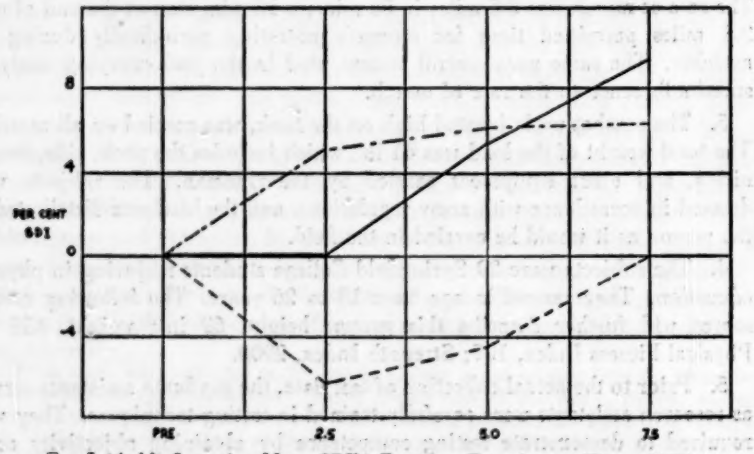


FIG. I. Ankle Inversion Mean SDI's, Footwear Experiment (Data: 1953).
——— = rubber insulated; — — — = combat; — . — . — = oxford

2. With the rubber insulated boot, the subjects did not experience fatigue of these muscles at 2.5 miles. However, fatigue increased for the balance of the march (almost a straight-line rise) to a strength loss of 9.5 per cent. The difference between the pre-march mean and the mean at 7.5 miles was 3.2 lb.; the t was 1.97, which is significant beyond the .05 level of confidence.

3. For the quarter-cut shoes, fatigue of the ankle inverter muscles was felt within the first 2.5 miles, a strength loss of 5.0 per cent. Little change was noted during the remainder of the march. The t for the difference between the pre- and post-march strength means was 2.11, indicating significance between the .05 and .01 levels.

ANKLE EVERTERS

The SDI's for the ankle everter muscles on the three footwear marches appear in Figure II. Comments on these results are:

1. Wearing the rubber insulated boot resulted in the least strength loss for these muscles; there was a slight increase in strength at 2.5 miles, at which point a steady decline followed for the balance of the march. At 7.5 miles, the SDI approached but did not reach the .05 level of confidence (t of 1.62).

2. With the combat boot, the decline in strength at 5.0 miles was significant at the .05 level; the t was 1.77.

3. The quarter-cut shoe caused by far the greatest strength loss for the ankle everter muscles at 7.5 miles. The SDI was 11.5, with a t of 3.72, which is significant beyond the .01 level.

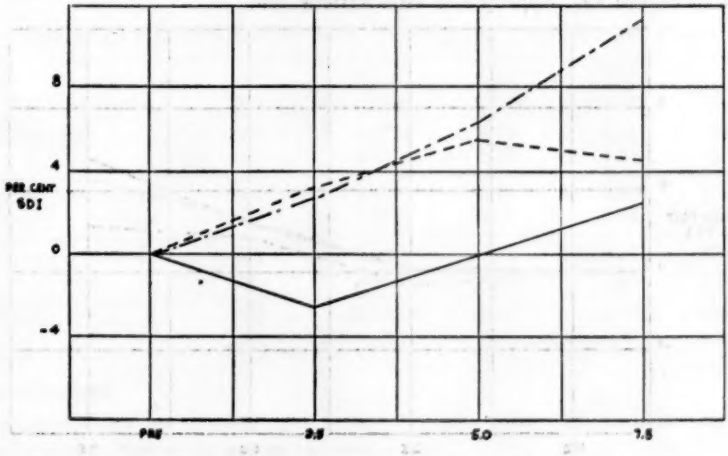


FIG. II. Ankle Eversion Mean SDI's, Footwear Experiment (Data: 1953).
——— = rubber insulated; — — — = combat; — · — · — = oxford

ANKLE DORSI FLEXORS

In Figure III, the SDI's for the ankle dorsi flexor muscles are presented. The strength of these muscles improved on all three marches. However, only one increase was statistically significant: the rubber insulated boot at 5.0 miles (t of 2.63; significant at the .01 level).

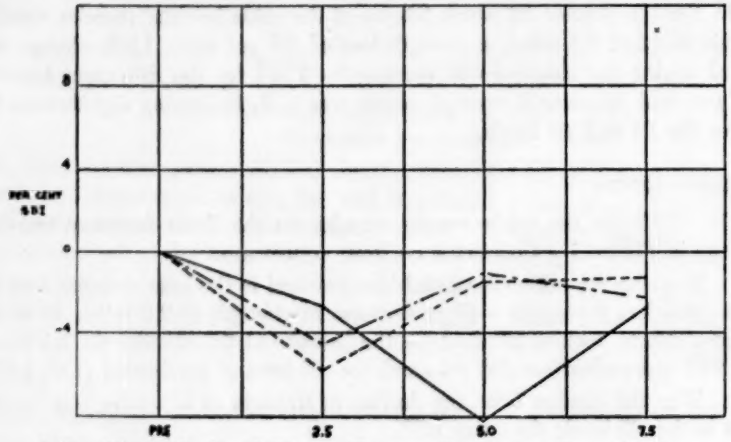


FIG. III. Ankle Dorsal Flexion, Footwear Experiment (Data: 1953).

——— = rubber insulated; — — — = combat; — · — · — =oxford

ANKLE PLANTAR FLEXORS

The SDI's obtained on the marches for the ankle plantar flexor muscles appear in Figure IV. Appropriate observations are:

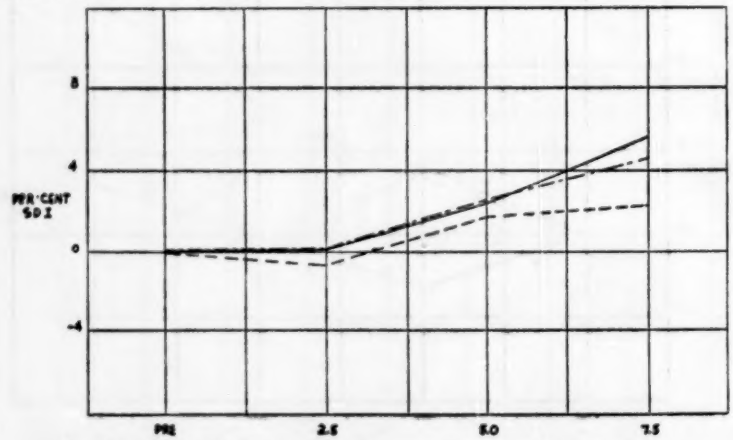


FIG. IV. Ankle Plantar Flexion, Footwear Experiment (Data: 1953).

——— = rubber insulated; — — — = combat; — · — · — = oxford

1. No stress on these muscles was noticeable at 2.5 miles for all three types of footwear. At 7.5 miles, however, significant mean strength losses were found with the rubber insulated boot and the quarter-cut shoe; the respective t's were 1.84 and 1.89, both being significant beyond the .05 level of confidence.

2. The loss in the strength of the ankle plantar muscles for the march in which the combat boot was worn was not statistically significant (t of 1.37). These results are not in agreement with those of the pack-carrying study, where strength losses of these muscles were significant for all six marches of the first series. However, the differences between the two series are within the range of chance probability.

A general summarization of strength decrements for the three marches wearing different army footwear appears in Table 1. Inasmuch as mean strength losses of the ankle dorsi flexor muscles were not found for any of the marches, these data are omitted from this table. Both the mean SDI's and the t's for the differences between pre- and post-march strength means are given for each muscle group on each march. The most significant figures on this table are the t's. The mean SDI's are affected by any unusual situations resulting in extreme scores, which could happen inadvertently during any march; such scores increase means disproportionately to the value of single cases. However, these scores also increase the variability (standard deviations) of the data, which, in turn, increase the standard errors of these measures. Statistically, then, low t's may be found with high means, as for the ankle everters on the march with the quarter-cut shoes. Thus, greater confidence can be placed in the t values.

TABLE 1

Summarization of Strength Decrement Indices for the Three Footwear Marches

Footwear March	Mean	Ankle Inverters	Ankle Everters	Ankle Plantar Flexors	Totals
Combat boot	SDI	0.00	5.39	3.21	8.60
	t	0.00	1.77	1.37	3.14
Rubber insulated boot	SDI	9.43	2.41	5.61	17.45
	t	1.97	1.62	1.84	5.43
Quarter-cut shoe	SDI	6.40	11.64	4.55	22.59
	t	2.11	3.72	1.89	7.72
Total	SDI	15.83	19.44	13.37	48.64
	t	4.08	7.11	5.10	16.29

Conclusions

Observations on the strength losses resulting from the various marches performed in this study are as follows:

1. Military marches with the combat boot resulted in the least fatigue of the muscles of the lower leg and ankle, as reflected in Strength Decrement

Indices of the muscle groups affected. The lowest total SDI (8.60) was obtained for this march. Only strength losses in one muscle group, the ankle everters, were statistically significant at the .05 level of confidence (t of 1.77).

2. The rubber insulated boot was the next favored of the army footwear studied. The total strength loss for this boot was approximately 75 per cent greater than for the march with the combat boot. All three muscle groups showed strength losses which were significant at the .05 level or nearly so.

3. Strength loss with the quarter-cut shoe was the greatest of the three marches. The total strength loss was nearly two and one-half times greater than for the combat boot march; and approximately 40 per cent greater than for the march with the rubber insulated boot. The strength losses for two muscle groups—ankle inverters and ankle plantar flexors—were significant beyond the .05 level of confidence; and the loss for one, the ankle everters, beyond the .01 level.

4. The muscle group most affected by proportionate strength losses for three footwear marches was the ankle everters. Statistical significance was found for the different marches as follows: rubber insulated boot, nearly at the .05 level; combat boot, slightly beond the .05 level; and quarter-cut shoes, beyond the .01 level.

REFERENCES

1. Clarke, H. Harrison, *A Manual: Cable-Tension Strength Tests.* Chicopee, Mass.: Brown-Murphy Co., 1953.
2. ———, Clayton T. Shay, and Donald K. Mathews, Strength decrements from carrying various packs on military marches, *Research Quarterly*, **26**: 253-265, 1955.
3. Garrett, Henry E., *Statistics in Education and Psychology*, 4th ed. New York: Longmans, Green and Co., 1953.

(Submitted 10/5/54)

Effect of Trained and Untrained Testers Upon the Administration of Grip Strength Tests

GEORGE F. COUSINS
Indiana University
Bloomington, Indiana

Abstract

In this study, scores obtained from two administrations of the test of right grip given to college freshman physical education majors by trained and untrained testers were compared.

Trained and untrained testers were equally effective in the administration of the test of right grip. No statistically significant differences were found between testers, between trials, or between the interaction of trials and testers.

IN THE DECEMBER, 1953 issue of the *Research Quarterly*, Mathews (1) reported the results of a study which compared scores when PFI tests were given to non-oriented testers with scores resulting when the same subjects were oriented and the testers experienced. Statistically significant differences were found in all Physical Index texts in favor of the experienced testers, and the conclusions indicated the necessity for trained examiners and pre-test practice sessions.

Because of the design of the experiment, obtained significant differences might have been the result of 1. the influence of the experience of the testers, 2. the effect of practice or learning, or 3. a combination of interaction of both experience and learning. The design of the experiment does not permit a precise interpretation of the resultant scores. The only justifiable conclusion that could be made is that a significant difference between means was found. To conclude that the learning factor influenced the more *valid* scores, and practice sessions should precede the actual testing program may not be supported by further research.

This particular design does not explain which of the above factors caused difference between the results. The influence or learning, or the influence of testers *and* practice should be investigated, as well as the influence of the training of testers. In order to further examine the effect of the training of testers and re-tester experience, the following related study was instigated.

Purpose

It is the purpose of this study to compare the scores obtained from the test of right grip given to college Freshman physical education majors by trained and untrained testers.

Procedure

The subjects utilized in this study were 40 Freshman physical education majors enrolled at Indiana University in a Freshman Major Activity Tech-

nique Course. Fifty subjects were initially selected at random from a total of 63 major students, and assigned at random to ten groups. Five students constituted one group. The ten testers used in this study were male students enrolled in the graduate class, Measurement and Evaluation.

Five of the testers selected at random were given detailed instructions concerning the administration of the test of right grip. These five individuals administered the test of right grip to unselected non-major Freshman students enrolled in the required physical education program. Using the tensiometer, each of these testers administered the test of right grip a minimum of 35 times. During this practice period, questions concerning proper administration were answered and a final training period was held. These five testers were then classified as being trained in the administration of the test of right grip.

The remaining five graduate testers, none of whom had previously administered the test of right grip, were then read the discussion of proper administration of the test of right grip as given by Clarke (2, *p. 158*). The instrument to be used in the study was shown and the administration of the test with the use of the tensiometer was discussed. No further training was given to this group, and these five testers were designated as untrained testers.

Each of the testers was then assigned at random to one of the ten groups of subjects.

The test of right grip was then administered twice to each group by the assigned tester. The initial test was given on a Tuesday and the final test was given on Thursday of the same week.

The instrument used was a tensiometer, model T5-6007-114 calibrated for a range from 0 to 200 pounds on $\frac{1}{16}$ in. size standard flexible aircraft cable. This tensiometer is guaranteed to read accurately within 5 per cent of the load, and with careful use closer accuracy could be obtained.

An attachment model T579 for use with the tensiometer originally developed by the Pacific Scientific Company in co-operation with Dr. Becktol of San Francisco makes it possible to test grip strength.[1] Figure I shows the tensiometer and the attachment.

The dial of the tensiometer does not directly indicate pounds of pressure, but must be transformed by means of a Riser Load Table. Obtained scores in this study were read directly from the instrument and were not converted to pounds of pressure.

Because of absences and injuries, one individual from each of five different groups did not take both tests. Of the five groups in which all subjects took both tests, one individual was dropped by random technique. This resulted in ten groups, four subjects to a group, for a total of 40 subjects. Five groups were administered the test by the designated trained testers and the remaining five groups by the untrained testers.

---

[1]This instrument and attachment are available and may be obtained from the Pacific Scientific Company, 1430 Grande Vista Avenue, Los Angeles 23, California.

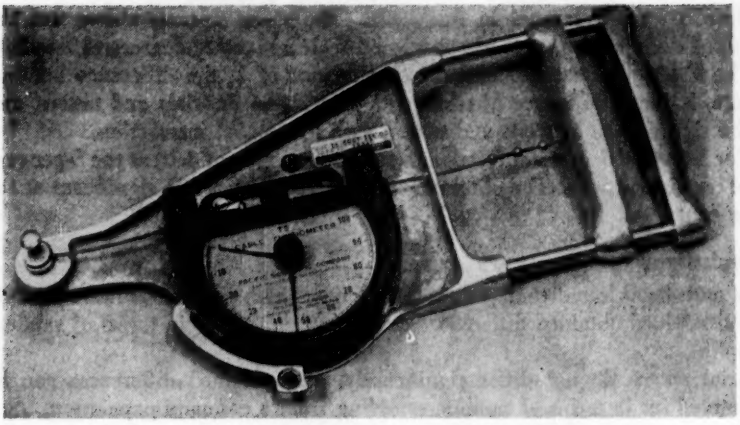


FIG. I. Tensiometer and attachment.

Method of Analysis

The experimental design used in this study is the analysis of variance applied to combined groups of subjects, each group having two trials. This analysis of repeated measurements on several independent groups is described by Edwards (3, *p. 288*).

Results

The obtained coefficient of correlation for the test of right grip between the first and second trials was .849±.04. This reliability coefficient compares favorably with that reported by Clarke (2, p. *172*).

TABLE 1

Analysis of Variance of Performance Scores of Right Grip of Two Groups of Subjects Tested by Trained and Untrained Testers with Two Trials for Each Group

Source of Variation	Sum of Squares	df	Mean Square	F
Between testers	221.1125	1	221.1125	1.5195
Between subjects in same group	5529.2750	38	145.5072	
Total between subjects	5750.3875		39	
Between trials	12.0125	1	12.0125	.9671
Interaction trials and testers	1.5125	1	1.5125	.1217
Interaction pooled subjects and trials	471.9750	38	12.4203	
Total within subjects	485.5000		40	
Grand Total	6235.8875		79	

Table 1 contains the results of the analysis. The various sums of squares obtained by division of the total sum of squares, and the associated degrees

of freedom are presented in this table. The mean squares shown are the result of dividing the sums of squares by their respective degrees of freedom.

The F test was used to find the significance of 1. the difference between testers, 2. the significance of trials with interaction of trials and testers, and 3. the significance of trials and testers with the pooled interaction.

For 1 and 38 degrees of freedom F requires a value of 4.10 at the 5-per-cent level for significance. The obtained variance ratios were not significant at the 5-per-cent level.

Conclusions

It may be inferred that the testers (trained and untrained) were equally effective, since obtained differences can be accounted for in terms of random sampling.

Trial means do not differ significantly, and obtained differences can be accounted for in terms of random sampling from a common population.

The interaction of trials and testers does not influence the obtained results.

These interpretations are made as a result of the above experiment utilizing the test of right grip strength, and cannot be made in respect to other testing situations.

REFERENCES

1. Mathews, Donald K., Comparison of Testers and Subjects in Administering Physical Fitness Index Tests. *Research Quarterly*, **24**: 442-445, 1953.
2. Clarke, H. Harrison, *Application of Measurement to Health and Physical Education.* New York: Prentice-Hall, Inc., 1950.
3. Edwards, Allen L., *Experimental Design in Psychological Research.* New York: Rinehart and Co., Inc., 1950.

(Submitted 10/12/54)

A Factor Analysis of Selected Wartime Fitness Tests

GEORGE F. COUSINS
Indiana University
Bloomington, Indiana

Abstract

In this study Thurstone's method of multiple factor analysis was utilized to factor the intercorrelations of 26 selected wartime fitness test items administered to 116 male Freshmen enrolled at Indiana University. Four factors isolated were identified in terms of the resultant correlations between the individual test items and the factors. Isolated factors were identified with reservation as: I Arm Extensor Endurance, II Power of Leg Extensors, III Power of Hip Extensors, and IV Power of Thigh Flexors. Low communalities of each of the test items indicated the presence of a specific factor or factors not isolated in the study.

THE TERM physical fitness has many connotations but no commonly accepted definition. This phrase has been consistently accepted throughout the field of physical education as one of its objectives. Lack of a clear, concise and acceptable definition of the term has caused many problems in the measurement of physical fitness.

The program of physical education was utilized during World War II to develop and improve the physical condition of the members of the Armed Forces. Outstanding leaders in athletic and physical education were selected to plan and conduct physical fitness programs for the various branches of the services. National, state, and local physical fitness offices were set up to advise and administer programs for civilians. As an outgrowth of these established programs, many tests of physical fitness were introduced by branches of the Armed Forces as well as by individuals in school and civilian programs. Although these tests were designated as physical fitness tests, referring to the common interpretation of the term, they are in reality tests of motor fitness.

Motor fitness is a quality which cannot be measured directly; however, it is measurable through the motor performance of the individual. There are many and varied types of motor performance, but the primary concern of wartime fitness is the large muscle type of movement over reasonable periods of time as indicated by selected wartime tests.

One of the difficulties of measurement in the area of fitness has been the lack of understanding of the degree of inter-relationships among the components of motor fitness. There is a need for the elimination of undesirable items and the establishment of an instrument to measure all common factors without overlapping and duplication. Selected physical tests supposedly measure the qualities necessary for an efficient fighting man in time of war. What are these qualities? What factors comprise the physical condition de-

sired by the Armed Forces? Through the factor analysis technique, it should be possible to answer these questions to some extent and in so doing improve methods and techniques of obtaining this quality of physical condition.

It is the purpose of this study to analyze a number of selected wartime physical fitness test items by Thurstone's method of factor analysis. Through the factor analysis technique and partial and multiple correlation, it should be possible: (a) to discover the primary factors or components of the selected wartime physical fitness tests; (b) to discover relationships of these factors and the relative importance of each factor; (c) to devise a battery of items which would best measure common factors basic to the tests; (d) to determine the validity of the devised battery of tests; (e) to develop regression equations for the prediction of each of the isolated factors; and (f) to develop regression equations for the best combination of all factors which would be considered as an index to measure all of the factors of wartime fitness tests with a high degree of validity.

Procedure

One hundred and seventy-four college Freshmen were used as subjects in this investigation, but only 116 of this number were used in the final analysis. All men who participated in the study were enrolled in the required physical education program. Records of individuals who failed to complete all tests were discarded. Reasons for failure to complete all tests were numerous: drop-outs, illness, and enlistment in the armed forces accounted for the majority of cases discarded. The mean age of the 116 subjects was 18.5 years, the mean height was 5 ft. 10 in., and the mean weight 152.4 pounds.

Even though many motivating forces were introduced, it was felt and observed that several students were not giving valid and reliable performances of their ability. The records of these individuals were not honored since the reliability of the items would be influenced by this type of behavior.

The testing program was carried out in nine class periods throughout the semester with the regular program of course content being followed as closely as possible. Because of this method of administration, rather than consecutive administration of the tests, the "guinea pig complex" was not observed in the attitude displayed by students for these tests.

The first class period was devoted to an explanation of the testing program and its purpose. The students were informed that the program was a part of the research being carried out by the department of physical education. Since all the subjects were already at or nearing the draft age, the performance of the tests would aid in the development of their physical condition. Also, since this testing was included as a part of the course work, final marking would be influenced by performance during testing. The subjects were informed that, while the final mark did not depend upon how well they performed on the various items, it might be influenced by the completion of all test items and the co-operation and attitude previous to and during testing periods.

A two-minute warm-up period, consisting of three continuous exercises, was administered before each testing period with exercises standardized throughout the testing period. The exercises consisted of (a) side straddle hop, (b) in-place running, and (c) 10-sec. squat thrust. All test items were performed by subjects dressed in regulation gymnasium uniform consisting of shoes, socks, shorts, and shirts.

Tests were administered with the aid of four graduate assistants majoring in physical education, and seven senior major assistants. All tests were administered on a competitive basis, which was an added stimulus for every individual to give his best effort. The purpose of each test was briefly explained prior to the actual testing, and where necessary, practice in each event was allowed. Results of each test were interpreted to the students in order to gain interest and to inspire effort.

The reliability of the test items used in this study was not obtained because of the large number of variables and the limitation of time for test administration. It was imperative that all tests be administered during the first semester since the individuals comprising the sample would be separated by the second semester.

Criteria for the selection of the test items were set up empirically by the writer. The criteria used for the selection of tests for this study were as follows: (a) They must be tests developed and used in the period of World War II. (b) They must be continuous test variables as apart from pass or fail items. (c) They must be items which could be administered to a large number of men using the available facilities at Indiana University. (d) They must be applicable to the age level of subjects. (This was included to eliminate fitness tests developed for elementary and junior high school students.) (e) They must be items which meet the commonly accepted criteria of validity, reliability, and objectivity.

Analysis of the literature pertaining to fitness tests resulted in the empirical selection of test items which best fitted the established criterion. The 31 variables selected for use in this study are as follows:

1. Sit-ups (fatigue)
2. Pull-ups (fatigue)
3. 300-yard Shuttle Run
4. 60-second Squat Thrust
5. Push-ups (fatigue)
6. Squat Jumps (fatigue)
7. 100-yard Pick-a-back
8. 20-second Squat Thrust
9. Running Broad Jump
10. 16-pound Shot Put
11. 75-yard Pick-a-back
12. Dodging Run
13. 6-second Run
14. Straddle Chins
15. Vertical Jump
16. Standing Broad Jump

17. Dips
18. 250-yard Shuttle
19. 2-minute Sit-ups
20. 20-foot Rope Climb
21. 100-yard Dash
22. Standing Hop-step-jump
23. 75-yard Dash
24. 150-yard Dash
25. 220-yard Dash
26. Hurdle Run
27. Medicine Ball Put
28. 3 Standing Broad Jumps
29. Leg Lifts
30. 300-yard Run
31. 200-yard Shuttle Run

Physical fitness batteries used as a basis for the selection of test items were as follows:

1. Army Air Force Test
2. Navy Standard Physical Fitness Test
3. Army Specialized Training Division Physical Efficiency Test
4. Army Physical Efficiency Test
5. Army Physical Efficiency Test, 1946 Revision
6. Indiana University Motor Fitness Indices
7. Larson's Dynamic Strength Test
8. U. S. Military Academy Physical Performance Test Batteries
9. U. S. Office of Education—Achievement Standards

Obtained raw scores of 31 physical tests administered to 116 male Freshmen students were transformed into T-scores. T-scores of each of the 31 items were added to form a composite score. The composite score, which was T-scored, was used as a criterion of physical fitness. A table of intercorrelations of all test variables, plus the correlations of all test variables with the criterion measure, was developed. International Business Machines were a merchanical aid in this procedure. Table 1 shows the inter-relationships of the test items. By inspection, this matrix revealed that several tests did not have high correlations with any other test. These tests: squat jumps, 20-sec. squat thrust, straddle chins, 20-ft. rope climb, and leg lifts were eliminated from the final analysis. These particular tests were rejected because the emergent factors are dependent upon the original correlation. Since the correlations of these tests were not high enough to change the results of the analysis, they were not included.

The analysis of the 26 remaining items required the computation of four sets of factor loadings and four sets of factor residuals. The computation of the factor loadings and the factor residuals were considered complete when no residual was found to be significant at the 1-per-cent level of confidence. The unrotated factor loadings are given in Table 2.

TABLE 2

Unrotated Factor Loadings

Item	K_1	K_2	K_3	K_4	h_k^2
Sit-ups	.4344	—.2422	—.3356	.2076	.4031
Pull-ups	.4543	—.1109	—.2172	—.0581	.2693
300-yd. Shuttle	.3705	.1358	—.2557	.1909	.2575
60-sec. Squat Thrust	.3437	—.3124	—.0255	.1101	.2283
Push-ups	.4871	—.3706	—.2208	—.3030	.5152
100-yd. Pick-a-back	.3819	—.2756	.4222	—.2661	.4708
Running Broad Jump	.6913	.1761	.1274	—.0713	.5303
16-lb. Shot Put	.6646	.1953	—.1634	—.3026	.5982
75-yd. Pick-a-back	.4454	—.2334	.3957	.2049	.4515
Dodging Run	.4608	.2620	—.1724	—.0841	.3177
6-sec. Run	.4971	—.1793	—.0118	.2361	.3351
Vertical Jump	.6406	.0735	.1923	.0156	.4531
Standing Broad Jump	.6399	.0465	.1993	.0142	.4516
Dips	.5234	—.2229	—.1595	—.3176	.4499
250-yd. Shuttle	.3463	.2988	—.1566	.1143	.2468
2-min. Sit-ups	.4062	—.3545	—.2456	.2350	.4062
100-yd. Dash	.4472	—.2747	.2713	.0919	.3575
Standing Hop-step-jump	.4923	.3627	.0917	—.0615	.3862
75-yd. Dash	.7094	—.0910	.2657	.1673	.6100
150-yd. Dash	.7116	.1309	.0762	.0856	.5366
220-yd. Dash	.6821	.2853	.0511	.1895	.5853
Hurdle Run	.4739	—.0621	—.1447	.1697	.2782
Medicine Ball Put	.6061	.2918	—.0730	—.2518	.5213
3 Standing Broad Jumps	.6863	.0741	.1984	.0226	.5163
300-yd. Run	.6097	.1498	.1635	.1102	.4329
200-yd. Shuttle	.6340	.2920	—.2599	.0969	.5642

K = Isolated Factor.

The next step in the analysis was the rotation of the factor loadings to make possible a meaningful interpretation of the factors. Rotated factor loadings are given in Table 3. The communalities of each test are shown by the symbol h_k^2. Inspection of Table 3 shows the negative correlations to be at a minimum. It shows no negative correlation is significant at the 1-per-cent level of confidence. The communalities of the majority of tests are comparable, and all tests show a relatively high uniqueness.

Isolated factors were then identified, and in order that the results of the study could be made meaningful and useful, equations were developed which could be used as predictive indices for the various isolated factors, as well

TABLE 3

Rotated Factor Loadings

Item	K_1'''	K_2'	K_3'	K_4'	h_k^2
Sit-ups	.2853	.1064	.1515	.5362	.4032
Pull-ups	.3614	.2179	.1736	.2472	.2693
300-yd. Shuttle	.1073	.3472	—.0235	.3535	.2576
60-sec. Squat Thrust	.1648	—.0063	.3365	.2966	.2284
Push-ups	.6221	.0452	.3207	.1524	.5151
100-yd. Pick-a-back	.1663	.0466	.6305	—.2082	.4707
Running Broad Jump	.1693	.5900	.3877	—.0565	.5303
16-lb. Shot Put	.4726	.5867	.1740	—.0183	.5982
75-yd. Pick-a-back	—.1250	.1203	.6283	.1632	.4515
Dodging Run	.2365	.5015	.0184	.0998	.3177
6-sec. Run	.0857	.1951	.3657	.3949	.3351
Vertical Jump	.0890	.4796	.4527	.1011	.4531
Standing Broad Jump	.0944	.4588	.4704	.1041	.4515
Dips	.5631	.1798	.3068	.0797	.4499
250-yd. Shuttle	.0417	.4533	—.0550	.1913	.2468
2-min. Sit-ups	.2418	.0036	.2504	.5338	.4061
100-yd. Dash	.0347	.0905	.5693	.1549	.3575
Standing Hop-step-jump	.0513	.5979	.1572	—.0365	.3862
75-yd. Dash	.0235	.4018	.6225	.2462	.6101
150-yd. Dash	.1151	5695	.3889	.2188	.5367
220-yd. Dash	—.0003	.6657	.2780	.2547	.5853
Hurdle Run	.1594	.2675	.2072	.3719	.2782
Medicine Ball Put	—.0015	.6202	.1508	.3375	.5213
3 Standing Broad Jumps	.0972	.5102	.4824	.1179	.5164
300-yd. Run	.0064	.5162	.3780	.1540	.4331
200-yd. Shuttle	.2179	.6388	.0452	.3265	.5642

K = Isolated Factor.

as for the combination of all factors. The rotated factor loadings and the correlation matrix were used as a basis for the developed equations. The Doolittle method of multiple correlation was employed in all instances for the purpose of obtaining the Beta weights, or relative weightings, of each item used in a particular index.

Discussion and Interpretation

Factor number one is identified as arm and shoulder girdle strength and co-ordination or arm extensor endurance because of the relatively high correlation with push-ups and dips. Significant correlations with pull-ups, 16-lb. shot put, sit-ups, and two-min. sit-ups are also shown. Pull-ups and the 16-lb.

shot put are generally considered to be tests in which this factor is present. The fact that the hands were clasped behind the head during the performance of sit-ups and two-minute sit-ups might lead to the conclusion that sit-up tests performed in this manner are measuring arm and shoulder girdle strength and co-ordination to some extent. Push-ups have the highest correlation (.6221) with this factor.

Identification of this factor is also possible through the zero correlation of certain test items with this factor. No significant correlations are shown with any of the running events. The power events such as the vertical jump, standing broad jump, and standing hop-step-jump all show insignificant relationships with this isolated factor. This would tend to support the identification of this factor since the majority of the test items which show a zero or insignificant relationship to the factor, are items which are commonly accepted as not having the strength and co-ordination of the arm and shoulder girdle muscles as a primary factor in the tests themselves.

Factor two is identified as strength or power of leg extensors because of the correlations with tests that are known to include strength as a factor. The relatively high correlations with events which have been designated as power items substantiates this identification. McCloy states that power is equal to force times velocity. If we interpret force to mean strength, and velocity to mean speed, the correlations of the power events show significant correlations with the strength factor and also significant correlations with factor three. The highest correlation with this factor is the 220-yd. dash. It is interesting to note that as the distance of the dash events is increased, excluding the 100-yd. dash, the correlation with the factor is increased.

Several of the resultant correlations of the various test items with this factor of leg extensor power support the identification of this factor. Sit-ups and two-min. sit-ups do not show significant relationship to the isolated factor, nor does the 60-sec. squat thrust which is generally considered a measure of agility. There is a rather large difference shown between the correlation of pull-ups with the factor when compared to the existant correlation of push-ups and this variable; however, the large correlation is not significant, and these tests appear to the primary measures of factor one.

Factor three is identified with reservation as a velocity factor because of the relatively high correlations with test variables in which speed is known to be a factor. Identification as power of hip extensors is plausible since known power items also show significant correlations with this factor. The 100-yd. pick-a-back has the highest correlation (.6305) with this factor; comparable correlations also exist between the 75-yd. pick-a-back (.6283), the 75-yd. dash (.6225), and this third factor.

It is interesting to note that the shuttle runs all show a zero correlation with this factor. It would seem that velocity should be a factor in those events; however, the relationships of the shuttle runs appear to be more highly related to factor two and also to the fourth isolated factor.

Significant relationships with the majority of test variables are shown with this factor. This appears plausible since the majority of tests have speed commonly accepted as a part of the particular test items.

Several zero or low relationships of test items (dodging run and 60-sec. squat thrust) are shown with this factor. These discrepancies are not meaningful and any attempt to rationalize the results would be invalid.

Factor four is identified as power of thigh flexors with the identification based on the correlations of the various tests of this factor. The highest correlations with this thigh flexor strength factor are sit-ups to fatigue (.5362) and two-min. sit-ups (.5338), and many variables which are known to have common use of the thigh flexor muscles also have significant correlations with this factor.

It would appear that this factor could be identified as abdominal strength; however, the literature appears to support the contention that sit-ups are more a measure of thigh or hip flexors than of abdominals. The Psoas Major, the Psoas Minor, the Iliacus, the Pectineus, and the Sartorius are the principle flexors of the hip or thigh. The abdominal muscles do come into action to some extent in the correct performance of sit-ups; however, they are not the principal muscles involved. If strengthening the abdominals is to be one of the major considerations of military tests, some other test must be employed which will more adequately measure abdominal strength.

Several factors are included in the battery of the 26 variables analyzed which are ordinarily thought of as tests of agility. Agility has been defined by McCloy as the ability to change very rapidly the direction of the body or parts of the body in space. The dodging run has a relatively high (.5015) correlation with the leg extensor factor, and no correlation with the factor of hip extensors. The relationship of this test variable to factor two seems plausible, since in this event the performer must stop and change direction quickly. The general strength of the individual would be a contributing asset for success in this event. It appears that there should be a significant correlation with factor three, but the writer has no explanation for the zero relationship.

The hurdle run, supposedly a measure of agility and a test similar in nature to the dodging run, is not subject, in this study, to the same interpretation. The correlations of this test and the factors show an insignificant relationship to factor one, but significant correlations with the remaining three factors. The highest correlation (.3719) of this test was with factor four.

The 60-sec. squat thrust or Burpee showed no significant relationship to either factor one or factor two. The highest correlation (.3365) was with the factor three. The communalities of these three tests of agility are relatively low, although significant. It appears that further analyses of tests of agility are necessary in order to discover a measure of the trait as defined previously.

Speed tests, as measured by the dashes and runs, were included in the analysis. As expected, the correlations of these tests with the factor three are

all significant. It is well, perhaps, to emphasize that for the most part as the length of the dash or run increases, the relation to the factor three decreases; and an increase in relation to the factor two is evidenced.

The 75- and 100-yd. pick-a-backs are commonly thought of as tests of speed and strength; however, the correlations of these events with factor two are very low. The only significant correlations of these tests to a factor are the relationships to factor three. The correlation of the 75-yd. pick-a-back is .6283, and that of the 100-yard pick-a-back is .6305, when related to the power of hip extensor factor. Thus, it is quite probable that the pick-a-back tests are measures of hip extensors more than measures of power of leg extensors.

The results of the 100-yd. dash are not in harmony with the interpretations of the 75-, 150-, and 220-yd. dashes with reference to the power of leg extensors factor. Although the 100-yd. dash has a high correlation (.5693) with the power of hip extensor factor, it shows no significant relationship to any of the three additional isolated factors. This conclusion is difficult to explain. It would seem that the resultant correlations would agree somewhat with the results of the other speed or dash events; however, this is not the case, and no explanation is presented.

The 6-sec. run shows a significant correlation with the factor three, but the highest relationship of this test is with factor four. These correlations are not appreciable, and since the communality is only .3351 it appears that there may be a specific factor in this test not common to any of the other test items.

The shuttle runs included in this study all show significant correlations with the factor two and the thigh flexor factor. It is commonly thought that these events would also correlate highly with the factor three; however, no correlation is evidenced in this study. These tests require the ability to use the large muscle groups of the legs to stop quickly and change direction. Speed between the turning points does not appear to be an important ability in these tests as does the ability to stop quickly and reverse direction of movement.

When the shuttle was removed from the 300-yd. event (shown by the 300-yd. run), the correlation with the factor two was increased. It is interesting to note that the correlation with the factor three was also appreciably increased.

Power was measured by a series of tests commonly accepted as meeting the force times velocity principle. It is evident that all of these tests—running broad jump, standing broad jump, vertical jump, standing hop-step-jump, and three standing broad jumps—are in close agreement. All show similar correlations with the factor two, as well as with the factor three. Apparently, then, all these tests are measures of power if we interpret strength as force and speed as velocity.

Sit-ups and two-min. sit-ups are commonly thought of as being pre-

dominantly measures of abdominal strength. Football coaches, wrestling coaches, and physical education instructors for some time have been using this event as the best method of increasing the abdominal strength of an individual. It is the belief of the investigator that sit-ups and variations of sit-ups are more indicative of the strength of the trunk flexors than of the abdominal muscles. The two sit-up events included in this study have the highest relationship to the thigh flexor factor. Significant correlations are also observed between these two events and factor one. This might be explained by the fact that the hands clasped behind the head appear to aid considerably in the performance of this test.

Pull-ups, push-ups, and dips are all generally considered to be measures of arm and shoulder girdle strength. Correlations of these three tests with factor one are high when compared to other test correlations with this factor. Significant correlations are also apparent with the factor three. This seems a reasonable expectation since the more rapidly these tests are performed, the better the performance in terms of number completed.

The 16-lb. shot put shows high correlations with the factor two. This result is in agreement with the general conception of the abilities involved in putting the shot.

The medicine ball put is generally considered a power item and shows a high relation (.6202) with the factor two. A significant correlation is also observed between the medicine ball put and the thigh flexor factor. The negative and low relationships with factors one and three are not meaningful, and no explanation is attempted. Although there are several instances where no explanation can be given relative to the correlations derived, the majority of the test and factor intercorrelations are capable of meaningful interpretation.

Regression equations in standard score form were developed for predicting each of the isolated factors. Additional predictive indices were developed for 1. combined factors, 2. test items having the highest communalities, 3. test items selected according to the principle of being relatively easily administered, and 4. test items selected on the basis of a high correlation with the criterion measure and a low intercorrelation between test items. All of the above equations for prediction are available from the author.

Conclusions

1. Of the 26 selected test items which are supposedly measures of strength, speed, agility, endurance, power, balance, flexibility, co-ordination, and cardio-respiratory endurance only four factors were identified. The low communalities of each of the test items indicates the presence of a specific factor or factors not isolated.

2. The four factors isolated were identified in terms of the resultant correlations between the individual tests and the factors. Isolated factors were identified as:

I Arm Extensor Endurance
II Power of Leg Extensors
III Power of Hip Extensors
IV Power of Thigh Flexors

Factors are identified with considerable reservation because of the low factor loadings and communalities.

3. Regression equations in standard score form are available for prediction in terms of isolated factors.

4. A predictive index for combined factors was developed and is available.

5. An index of test items having the highest communalities was developed.

6. The recommended four-item battery selected on the basis of a high correlation with the criterion measure and a low intercorrelation between test items showed a multiple R of .9040. The equation is:

$$X_c = -6.7252 \text{ (75-yard dash)} + .0797 \text{ (16-pound shot put)} + .1871 \text{ (3 standing broad jumps)} + .5601 \text{ (dips)} + 38.0657$$

7. The results of this study are not in harmony with the measurements of agility as defined. Strength appeared to be the most significant factor in the successful performance of this trait.

8. Dashes included in this analysis showed high relationships with the isolated factor of hip extensors.

9. Recommended test batteries are not in agreement with the batteries developed and used during the period of World War II.

10. There is a need for further study in the area of motor fitness in order to determine what other factors, if any, these test items are measuring.

REFERENCES

1. Appleton, Lloyd O., *The Practicability of Standardized Procedures for Physical Performance Tests*, Office of Physical Education, U. S. Military Academy, West Point, New York, 1949, mimeo., n.p.
2. Bookwalter, Karl W., A survey of factor analysis studies in physical education, *The Physical Educator*, **2**: 209-212, 1942.
3. Brogden, Hubert E., and Harry H. Harman, *An Analysis of Factors in Physical Proficiency*, Personnel Research Section, Adjutant General's Office, Department of the Army, Washington, D. C., 1948, 6 pp., ditto.
4. Carpenter, Aileen, An analysis of the relationships of the factors of velocity, strength and dead weight to athletic performance, *Research Quarterly*, **12**: 34-39, 1941.
5. ———, A critical study of the factors determining effective strength tests for women, *Research Quarterly*, **3**: 3-32, 1938.
6. ———, The Differential Measurement of Speed in Primary School Children, *Child Development*, **12**: 1-7, 1941.
7. Cattell, Raymond B., Parallel proportional profiles and other principles for determining the choice of factors by rotation, *Psychometrika*, **9**: 267-283, 1944.
8. Coleman, James W., The differential measurement of the speed factor in large muscle activities, *Research Quarterly*, **8**: 123-30, 1937.
9. ———, Pure speed as a positive factor in some track and field events, *Research Quarterly*, **11**: 47-53, 1940.

10. Cook, Ellsworth B., and Robert J. Wherry, A statistical evaluation of physical fitness tests, *Research Quarterly*, **21**: 94-111, 1950.
11. Cureton, Thomas Kirk, *Analysis of Some Basic Factors in Motor Fitness*, Urbana, Ill., 1942, n.p., mimeo.
12. ———, *Physical Fitness Appraisal and Guidance*, St. Louis: C. V. Mosby Co., 1947, 566 pp.
13. ———, What is physical fitness? *J. Health, Physical Education and Recreation*, **16**: 1: 111-112, 1945.
14. ———, *et al.*, *Endurance of Young Men, Analysis of Endurance Exercises and Methods of Evaluating Motor Fitness*, Society for Research in Child Development, National Research Council, Washington, D. C., 1945, 284 pp.
15. Darling, Robert C., *et al.*, Physical fitness—report of the subcommittee of the Baruch Committee on Physical Medicine, *Journal of the American Medical Association*, **136**: 11: 764-767, 1948.
16. Guilford, J. P., *Psychometric Methods*, New York: McGraw-Hill Book Co., 1936, 566 pp.
17. Hall, D. M., and J. R. Wittenborn, Motor fitness tests for farm boys, *Research Quarterly*, **13**: 432-443, 1942.
18. Holzinger, Karl J., and Harry H. Harman, *Factor Analysis*, Chicago: University of Chicago Press, 1941, 417 pp.
19. Hotelling, H., Analysis of a complex of statistical variables into principal components, *J. Educational Psychology*, **24**: 417-441, 1933.
20. Kelley, Truman L., *Crossroads in the Mind of Man*, Stanford, Calif.: Stanford University Press, 1928, 238 pp.
21. Larson, Leonard A., A factor analysis of motor ability variables and tests, with tests for college men, *Research Quarterly*, **12**: 499-517, 1941.
22. Latham, Darrell E., *Factor Analysis of the Illinois Motor Fitness Screen Test*, Unpublished Master's thesis, University of Illinois, Urbana, 1940, n.p.
23. McCloy, Charles Harold, A preliminary study of factors in motor educability, *Research Quarterly*, **11**: 28-39, 1940.
24. McCloy, Emma, Factor analysis methods in the measurement of physical abilities, *Research Quarterly*, suppl., **6**: 114-121, 1935.
25. Roggen, Arthur, A study of the relationships between the general factor in events primarily depending on strength, and the general factor in four broad jumps, using Spearman's method of analysis for G, *Research Quarterly*, suppl. **6**: 122-127, 1935.
26. Sigerseth, Peter O., *An Attempt to Isolate and Identify Fundamental Factors that Underlie Motor Fitness as Measured by Large Muscle Performance*, unpublished doctoral dissertation, University of Oregon, 1944, 71 pp.
27. Sills, Frank D., A factor analysis of somatotypes and their relationship to achievement in motor skills, *Research Quarterly*, **21**: 425-427, 1950.
28. Thomson, Godfrey H., *The Factorial Analysis of Human Ability*, Houghton Mifflin Co., New York, 1939, 326 pp.
29. Thurstone, Leon L., Current issues in factor analysis, *Psychological Bulletin*, **4**: 189, 1940.
30. ———, *Multiple Factor Analysis*, Chicago: University of Chicago Press, 1947, 535 pp.
31. ———, *The Theory of Multiple Factors*, Ann Arbor, Mich.: Edwards Bros., 1933, 65 pp.
32. ———, *The Vectors of the Mind*, Chicago: University of Chicago Press, 1936, 266 pp.
33. Wolfe, Dael Lee, *Factor Analysis to 1940*, Chicago: University of Chicago Press, 1940, 69 pp.

(Submitted 10/12/54)

Steady-Pace vs. Variable Speed in High School 220-Yard Run

FRED KRONSBEIN
University of California[1]
Berkeley, California

Abstract

Freshman high school boys were given three weeks of group practice in their physical education class, in running 80 and 160 yd. at a steady pace. A timing gong was sounded at the instant the runner was supposed to be crossing each of a series of 20-yd. marks. During the third week, individual time trials were made in a 220-yd. run without the gong. Two runs were of the steady-pace pattern and a third was made at "all-out" speed throughout to secure a variable-pace high-decrement pattern. Time trials were repeated during a fourth week. The steady-pace run average 0.6 sec. faster ($t = 3.85$) than the all-out run, and the boys reported less fatigue.

IT HAS BEEN KNOWN for many years that the oxygen requirement of running increases very rapidly as the velocity becomes greater. Physiologists have recommended the even pace as more economical of energy. Nevertheless, most coaches and athletes advocate and use a variable-pace design for running the longer sprints and middle distances.

Recommended Patterns

For the 220-yd. race, good runners should sprint the first 75 yd., float 75 yd. (while the runner slows his speed slightly), and then drive the last 70 yd., according to Gauthier and Haney (4). This opinion corresponds to Cromwell's: "If they will learn how to run their best 200 yds. in advance, knowing when to ease off from the opening burst of speed to the coast, and when to go from that to the final gather, they are most likely to win" (2, *p. 58*).

Doherty feels that in the 220, the problem is to achieve maximum velocity and then maintain it as long as possible. "The speed should very gradually but consistently fall away after the 60 yard mark through to the finish." He shows a typical all-out deceleration curve as an example (3, *p. 74*). In both the 440 and 220, he thinks that there should be a gradual and natural deceleration to make sure that the runner crosses the tape completely exhausted. He does not approve of the sprint-float-drive pattern recommended by Cromwell and others who advocate that system (1, 2, 4).

Murphy (9, *p. 42*) as early as 1914 thought that in running the 220-yd. dash the runner must learn to "swing through this race at almost one's best speed for the first 180 yd. without actually trying every step of the way as in the 100-yd. dash." In other words, he advocates a relatively even-pace type of run with less variation in speed than in the variable-pace patterns advocated by coaching textbooks.

---

[1] From the Research Laboratory of the Department of Physical Education. The present address of the writer is El Camino High School, Sacramento, California. Appreciation is expressed to Dr. Franklin M. Henry for advice during the course of the investigation.

Coaches who advocate the even-pace race have the support of the physiologists who have studied the problem of energy cost and speed of movement. This topic, particularly as it applies to running, has recently been reviewed by Henry (8). "It can be said with confidence that insofar as the physiological limit is involved in setting records, a steady pace will result in faster time for the 220 and 440 as well as the mile and two mile runs" (7).

Calculation of Advantage. Time-velocity equations which describe the position and speed of the runner at any time in either the steady (even) pace or the all-out pace have been developed (8). By using known oxygen requirement data and these equations, it is possible to estimate the amount of oxygen required for a run of a certain speed and velocity profile. By equating the metabolic costs of the two types of race, the relative efficiency of one type of race over the other has been computed (8).

No one seems to have investigated the comparative performance of runners in actually using the different types of velocity patterns under controlled conditions.

Purpose of Study

It is the purpose of this study to test the relative effectiveness of the steady-paced type of run compared with the variable-pace or high-decrement run, and to determine if it is feasible and practical to train high school students to run at a steady pace.

Steady pace is defined as the maintaining, after acceleration, of a constant velocity throughout the entire run. However, to insure that the runner has put forth full effort, there should be a slight dropping off in velocity during the very last part. In the variable pace, the runner will reach his maximum velocity in the early part of the run, and from that point on his velocity will decrease progressively.

Method

Subjects. Thirty-nine boys, who were volunteers from the regular Freshman physical education classes of the El Camino High School at Sacramento, California, were used in the study. None of them was experienced in running.

Apparatus. The set-up for the experiment was a modification of the one suggested by Henry (7). The starting signal for each run was given by a clap-board. The gong arrangement for the steady-pace training consisted of wooden tripod 5 ft. high with a steel bar, 1 in. in diameter and 18 in. long, hanging by a line from the center of the tripod. This hanging bar was struck vigorously in cadence with a metronome set according to the Henry tables (7). Tenth-second stop watches were used for time checks at 70, 130, 190, and 220 yd. for the time trials. The rates of these watches were all checked against the "standard" watch, and corrections made as required. The 220-yd. straightaway past the bleachers was used as a track, and the gong was set up in the bleachers approximately 80 yd. from the starting line. The check marks on the track for the training consisted of high hurdles placed on the

curb at 20-yd. intervals with the exception of the first one which was placed at 21 yd.

Procedure. The experiment began April 19 and lasted for five weeks. The subjects were told that the study would consist of a comparison of two ways of running. They did not know and were not told which type might be expected to give the best results. They were reminded several times during the course of the study of the importance of doing their best to run the way they were asked and to put forth every effort at all times. They were instructed in the start, running form, and the necessity of a suitable warm-up every day before the run. The medium-stance start, which has been found to be the best for most runners (5, 7) was used.

Training and Practice. First, the subjects ran 220 yards as fast as they could and the finish times were recorded. With these known times, the subjects were grouped to expedite their training for the steady-pace pattern. For each group, a metronome rate was chosen to give a total time just under their unit group's previous total time. No training group had over six subjects running at one time, and several groups had only one or two in them. The subjects were instructed to pull off of the track if they were not within one stride of the 20-yd. check marks at the sound of the gong.

During the course of the experiment the subjects were shifted from one group to another as their total times warranted, or if they thought they could run at a faster rate. After the first classification run, the subjects trained in their steady-pace groups for nine days (two calendar weeks). Each day they made two practice runs, the first a short one of 80 yd. and then a longer one of 160 yd. It had been noticed that the subjects had trouble keeping the first 40 yd. slow enough for the proper steady-pace pattern, so the 80-yd. distance was used to give additional practice on the first part of the run.

After the first nine practice sessions, individual time trials for the steady-pace run, without the gong, were made on two consecutive days with time checks at the 70-, 130-, 190,- and 220-yd. markers. The following day, each subject ran individually by the high decrement pattern (time checks were taken again at the designated spots). The experimenter and the timers gave verbal encouragement at each check point, as had also been done at the 190- and 220-yd. stations of the steady-pace runs. The subjects then practiced on the steady-pace run again for three days, and the 220-yd. individual time trial runs were repeated on two consecutive days. On the following day, the subjects each made a high decrement run of 220 yd. The next day (Friday) and on Monday and Tuesday of the following week, there was 80- and 160-yd. practice on the steady pace. On Wednesday and Thursday, two steady-pace 220-yd. runs were timed, and on Friday the last run (by the high decrement pattern) was made.

Results and Discussion

The scores of the last three time trials were examined statistically because it was felt that the subjects' performances were fairly stable by then, and that

there would no longer be much of a training or practice effect within a few consecutive runs. The very small difference between the two steady-pace runs seems to confirm this expectation; it was 0.03 seconds and not statistically significant ($t = 0.15$). However, the high decrement run was considerably slower, to the extent of 0.60 seconds. This difference is statistically significant ($t = 3.85$).[2] It is more clear-cut when the two steady-pace runs are averaged to secure greater reliability. The average of these, compared with the high-decrement run, gives a difference of 0.53 seconds (about the same as before) but the t ratio is larger ($t = 4.31$).

In order to balance out the possible training effect that might be present, we can look at the 220-yd. time trials for the last two weeks of the study. There were two steady-pace runs in the fourth week and two in the last week. In each case, the pair of steady-pace runs was followed by one high-decrement run. This gives us the average of four even-pace runs to compare with the average of two all-out runs (Table 1). This comparison shows the mean steady-pace was run 0.48 sec. faster than the mean high-decrement run, with the t ratio = 3.51 which is highly significant statistically.

TABLE 1

Time Trials Using Two Patterns of Pacing

Time	Pattern		
	Steady	Steady	High Decrement
Final week	30.68	30.65	31.26 sec
Previous week	30.92	31.44	31.54 sec
Average		30.92	31.40 sec

In Figure I it can be seen that the steady-pace run was not quite done in the theoretically perfect constant velocity pattern. However, the profile was definitely "flatter" than in the all-out or high-decrement type of run.

We do not know exactly how irregularly these runs were made. Assuming that the range of average velocities is the difference between the highest and lowest means as measured at the first timing station and the last or next to last (whichever was slowest), the greatest variation from the average velocity of the run was 5.7 per cent for the steady-pace run and 14.0 per cent for the

[2] The standard error of the difference, 0.167 sec., includes sampling error and timing error. While it was not determined, the mean timing error was probably small. Cureton and Coe (*Research Quarterly* 4: 94-109, 1933) give data that show the variable error of a single timer for a single run averages about 0.13 sec. Using this figure, the mean timing error for the difference between two runs would be 0.02 sec., which is negligible compared with the errors that arise from sampling and intra-individual variation. For a discussion of this problem by Henry *et al.*, see p. 465 of Scott (ed.), *Research Methods Applied to Health, Physical Education and Recreation*, Washington: Am. Assn. for Hlth., Phys. Educ. and Rec., 1949.

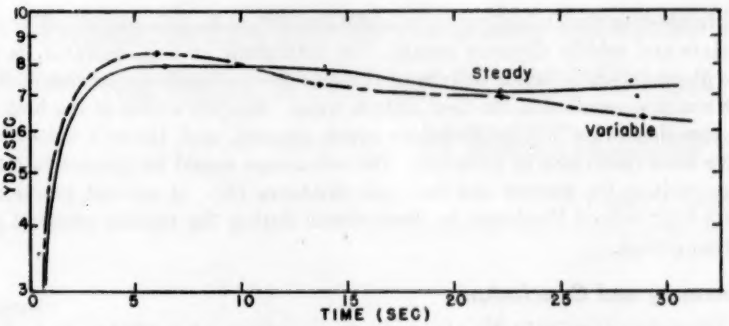


FIGURE I. Observed Mean Velocities During the Final Week.

high decrement run. Assuming that the area of discrepancy between the perfect steady pace and the run as made is triangular and the altitude at half its area approximates the mean discrepancy in velocity, these variations would be multiplied by 0.707 and become 4.02 per cent and 9.9 per cent.

From Henry's tables and curves (8), based on physiological efficiency, the steady-pace run would therefore have been done 0.64 per cent slower than a theoretically perfect even-pace, and the high decrement run would have been 2.75 per cent slower. In this experiment, therefore, the advantage of the relatively steady-pace run, when compared with the high-decrement run, should be 2.11 per cent. The observed time difference did significantly favor the steady-pace type of run and the observed advantage was 1.96 per cent in the time trials of the final week. Under the circumstances, the agreement between the theoretical and observed advantage is very high.

The motivation was probably better in the all-out high-decrement run than in the steady-pace run. In the latter, verbal encouragement could only be given near the end of the run, while in the all-out run the timers at each check point were able to urge the runner to go faster. The experimenter had the general impression that the subjects were more nearly exhausted in the high-decrement run. Verbal reports collected from the subjects, while not very well controlled or dependable, agree with this impression.

More practice was given in the steady-pace pattern than in the other. The reason was that judgment of pace had to be developed in order to make an unguided even-pace run. In the high-decrement run, judgment of pace is not required, as the subject runs as fast as possible all the way down the track. Other than this, there were no special skills developed that could favor one race pattern over the other one. With more practice, the boys would probably have come closer to the exact even pace. Older boys could be expected to learn the pacing more rapidly. An automatically starting, electrically amplified metronome would be very useful.

It is of interest that the steady-pace run in this experiment (which was intended to be closer to the "flat" profile than was actually accomplished) was

fairly close to the Doherty recommendation of "gradual deceleration" for the sprints and middle distance events. The important finding, however, is that the physiologically better pattern of the two profiles studied experimentally is the one that resulted in the best 220-yd. times. Murphy's idea of the best way to run this event (9) is therefore given support, and Henry's calculations have been confirmed in principle. The advantage would be greater in longer runs such as the quarter and half mile distances (8). It was not practical to train high school Freshmen in these events during the regular physical education period.

Summary and Conclusions

Thirty-nine male high school freshmen, inexperienced in track, were trained to run at a relatively steady pace for 220 yd. The purpose was to see if it was practical to teach them even-pace running, and to discover if this pattern of running was faster than the variable-pace pattern.

The steady-pace practice consisted of passing 20-yd. markers each time a pacing gong was sounded. The training periods were made up of short practice runs for a period of two weeks, followed by two weeks in which the subject on different days ran two time trials at an even-pace and one according to the high-decrement or "all-out" pattern. In the time trials, stop watches were used at the 70-, 130-, 190-, and 220-yd. markers in order to check the velocity patterns.

It was found that the 220-yd. speed in the steady-pace pattern was significantly faster than in the variable-pace pattern. The amount of advantage (2%) agreed well with the theoretical advantage calculated by Henry.

While three weeks of practice was insufficient to teach the boys to slow down enough in the first half of the run to achieve a true steady pace, they did come close enough to the desired pattern to show its advantage over a more variable velocity pattern.

REFERENCES

1. Bresnahan, G. T. and W. W. Tuttle. *Track and Field Athletics.* St. Louis: C. V. Mosby Co., 1947.
2. Cromwell, D. B. *Championship Technique in Track and Field.* New York: Whittlesey House, 1941.
3. Doherty, J. K. *Modern Track and Field.* New York: Prentice Hall, Inc., 1953.
4. Gauthier, G. and G. Haney. *Fundamentals of Track and Field Athletics.* Minneapolis: Burgess Publishing Co., 1951.
5. Henry, F. M. Force-time characteristics of the sprint start. *Research Quarterly,* **23**: 301-317, 1952.
6. Henry, F. M. and others, Individual differences in oxygen metabolism of work at two speeds of movement. *Research Quarterly,* **22**: 324-333, 1951.
7. Henry, F. M. Research in sprint running. *Athletic J.,* **32**: 32-34, 1952.
8. Henry, F. M., Time-velocity equations and oxygen requirements of "all-out" and "steady-pace" running. *Research Quarterly,* **25**: 164-177, 1953.
9. Murphy, M. C. *Athletic Training.* New York: Charles Scribner's Sons, 1914.

(Submitted 7/28/54)

Influences of Physical Education Activities on Motor Ability and Physical Fitness of Male Freshmen[1]

CARL W. LANDISS
Agricultural and Mechanical College of Texas
College Station, Texas

Abstract

This article presents an analysis of the effect of eight physical education activities on physical fitness and motor ability scores of male college Freshmen. The mean gain was computed for each group on both tests and individual test items. The mean gain of each group was compared with each of the other groups and significance of differences determined. The results seemed to indicate that of the activities studied tumbling-gymnastics and conditioning best developed those abilities measured by the physical fitness test; while tumbling-gymnastics and wrestling best seemed to develop those abilities measured by the motor ability test.

AMONG THE RECOGNIZED objectives of physical education are the objectives of developing physical fitness or motor fitness and motor ability. While considerable research has been done to establish tests to measure these abilities, relatively little experimental data are available concerning the contribution of a given activity to a specific objective. The value of an investigation of this type lies in the light it may shed on which activities best develop certain desired objectives within the participant.

The objectives considered in this study are limited to physical or motor fitness and motor ability as measured by the tests selected. The relative value of developing skills in certain activities, e.g., ability to swim, over-increasing physical fitness or motor ability was not considered in this research.

Statement of the Problem

This study was designed to compare eight selected physical education activities in their development of "physical fitness" and "motor ability" in the students participating in these activities. The criteria for determining these two factors of physical fitness and motor ability were the Physical Fitness Test (1) and the Larson Test of Motor Ability (5).

Related Studies

In 1943, Wilbur (7), in comparing two types of physical education programs, sports method versus formalized apparatus method, in terms of physical fitness test results, found:

1. The sports method evidenced significant superiority (total test) over

[1] This study was made in partial fulfillment of the requirements for the degree of Doctor of Education in the school of Physical Education, Pennsylvania State College, 1951.

apparatus method for the last one-half of the experiment. The sports method was also superior in developing arm and shoulder-girdle strength (chins and dips), and co-ordination, agility, and control (bar-snap).

2. The two methods were equal in improving speed of legs (dodging run), jumping or leg strength (jump and reach), arm and shoulder-girdle co-ordination (baseball throw), and endurance (one-quarter mile run).
3. The formalized apparatus method did not show significant superiority to the sports method on any of the tests used.

In a study of the influence of selected physical education activities on endurance test scores, Fordham (4) and Berrafato (2) in concurrent but separate studies found that students taking the basic conditioning course made the greatest improvement on the endurance test. Mean gains for each of the activities were: basic conditioning 19.35; apparatus 12.84; boxing 11.29; weight lifting 11.06; intramural athletics 9.40; individual tumbling 8.05; wrestling 6.50; volleyball 5.09; and badminton 3.07. The mean gain is indicated in terms of standard scores.

Activities Used

The activities selected for this investigation were elementary courses in swimming, boxing, weight training, tennis, wrestling, volleyball, tumbling-gymnastics, and a basic conditioning course. The conditioning course consisted of calisthenics, running, pull-ups, and grass drills. The tumbling-gymnastics course combined both mat tumbling and apparatus work. The classes met three times a week for one hour for instruction and participation in the designated activity. The primary objective for each course except the basic conditioning course was to develop playing or performance skill in that specific area. All classes were taught by experienced full-time college instructors.

Measures Used

The test used as the criterion measure to determine motor ability was the Larson Test of Motor Ability (5) consisting of four test items: 1. baseball throw for distance; 2. pull-ups; 3. vertical jump; and 4. bar-snap. The test used as the criterion measure of physical fitness was the Physical Fitness Test, which was a slight modification of the Army Air Forces PFR test (1). The Physical Fitness Test consisted of three test items: 1. 300-yard shuttle run; 2. pull-ups; and 3. sit-ups.

The term "motor fitness" might be more appropriately used than the term "physical fitness" which usually denotes "fitness for a given task."

Procedure

The study was conducted during the fall semester of the academic year 1950-51, at the Agricultural and Mechanical College of Texas. The subjects for the experiment were first-semester Freshmen at this institution. All incoming Freshmen (N = 1031) were tested by the Physical Fitness Test. Stu-

dents were matched by means of the physical fitness test and age in eight parallel sections. A third factor was used in the selection of the swimming group; namely, all students assigned to this group had failed a rather rigid swimming test.

An analysis for significance of difference on the group selection factors—(1) physical fitness test, and (2) age—showed that there were no significant differences between any of the eight groups on these two measures. All groups were tested for motor ability by the Larson Test of Motor Ability. A C.R. analysis of initial scores on the Larson test indicated a significant difference at the 5-per-cent level of confidence between two groups—weight training and conditioning, the weight-training group having the better scores. No other significant differences were noted. Each group, except the swimming group, was then arbitrarily assigned to one of the selected activities in which they participated and received instruction for one semester. At the end of the semester all groups were again tested using the same two measures. To reduce outside variables in physical education activities, all students were eliminated from the study who had participated in intramural activities more than three hours during the semester. Thus, the final sample was reduced to 432 students.

PHYSICAL FITNESS TEST

At the beginning of the experiment no significant differences existed between the eight groups on physical fitness scores. Seven of the eight experimental groups (Table 1) showed a significant improvement in performance on the physical fitness test.[2] The swimming group was the only group which did not improve significantly in abilities measured by this test.

A comparison of the mean gain of each group with the mean gain of each of the other groups is shown in Table 2. The t-test for the significance of differences showed that the tumbling-gymnastics and conditioning groups were superior to each of the other activities in developing physical fitness. There was no significant difference in these two activities when compared with each other. The test items used to determine physical fitness were practiced by the conditioning group as a part of the regular class period and it is quite possible that practice in the specific test items could invalidate the test for that group. The tumbling-gymnastics group was significantly superior to the tennis, volleyball, swimming, boxing, and weight-training groups. Condi-

[2] For quick reference in determining significance of t, the following figures may be used:

Even-numbered tables
- $t = 2.63$ is significant at the 1-per-cent level of confidence
- $t = 2.36$ is significant at the 2-per-cent level of confidence
- $t = 1.98$ is significant at the 5-per-cent level of confidence

Odd-numbered tables
- $t = 2.68$ is significant at the 1-per-cent level of confidence
- $t = 2.40$ is significant at the 2-per-cent level of confidence
- $t = 2.01$ is significant at the 5-per-cent level of confidence

TABLE 1

Mean Initial Scores, Gains, and Critical Ratios for Physical Fitness Test and Larson Motor Ability Test

Activity	N	Physical Fitness Test				Larson Motor Ability Test			
		Mean	S.D.	Gain	C.R.	Mean	S.D.	Gain	C.R.
Wrestling	58	29.50	3.16	3.97	3.61	594.31	46.01	45.86	5.47
Conditioning	50	29.56	3.88	5.86	7.10	588.60	44.53	27.00	3.18
Tennis	51	30.16	3.57	2.29	2.53	605.00	54.12	22.35	2.06
Volleyball	63	29.60	3.71	3.57	4.46	596.75	54.38	24.92	2.72
Swimming	52	28.85	3.19	0.63	0.89	597.31	50.08	12.50	1.25
Weight training	47	29.75	3.53	4.09	5.10	609.47	43.01	26.66	2.63
Tumbling-gymnastics	60	29.88	3.49	5.80	4.98	598.17	49.82	43.66	4.97
Boxing	51	29.94	3.56	3.12	4.79	601.47	46.73	13.33	1.37

TABLE 2

Critical Ratios of Differences in Mean Gains[1]

Activity	Wrestling	Conditioning	Tennis	Volleyball	Swimming	Weight Training	Tumbling-gymnastics	Boxing
Wrestling	.....	2.03	2.13	0.55	4.34	0.15	2.47	0.98
Conditioning	3.17	.....	4.17	2.69	5.94	1.90	0.07	2.82
Tennis	4.12	0.73	.....	1.75	2.16	2.20	4.81	0.95
Volleyball	3.66	0.33	0.42	.....	4.14	0.68	3.33	0.56
Swimming	5.29	2.09	1.47	1.85	.....	4.33	7.28	2.93
Weight training	3.07	0.05	0.65	0.26	1.97	.....	2.22	1.09
Tumbling-gymnastics	0.45	2.95	3.96	3.48	5.18	2.86	.....	3.27
Boxing	6.45	2.36	1.63	2.08	0.13	2.18	6.49	.....

[1] Figures above the dotted line in each column give ratios for the Physical Fitness Test; below, for the Larson Motor Ability Test.

tioning showed a significant superiority over tennis, swimming and boxing, volleyball and wrestling.

The data indicated that there was no significant difference in weight-training, wrestling, volleyball, and boxing in developing physical fitness in students who participated in those activities. All groups evidenced sufficiently superior to the swimming group so as to be considered significant at a level of confidence of at least 5 per cent.

LARSON TEST OF MOTOR ABILITY

As previously noted, the C.R. between the eight groups on the initial test showed two groups, conditioning and weight training, having a difference considered significant at the 5-per-cent level of confidence.

In analyzing the mean gain between the initial test and post-test for each group (Table 1), it was found that four activity groups—wrestling, conditioning, volleyball, and tumbling-gymnastics—made gains that were considered significant at the 1-per-cent level of confidence. The weight-training and tennis groups made gains that were significant at the 2-per-cent and 5-per-cent levels respectively. The swimming and boxing groups failed to make gains that were statistically significant.

The t-test for significance of differences in mean gains (Table 2) showed that wrestling and tumbling-gymnastics were superior to each of the other six activities in developing motor ability as measured by the Larson Test to a degree that was considered significant at the one per cent level of confidence. There was no significant difference when wrestling and tumbling-gymnastics were compared. Conditioning and weight training both developed motor ability as measured by the test used to a significantly greater extent than did swimming and boxing; while volleyball was superior to boxing. Tennis and volleyball seemed to be approximately equal in the development of motor ability in students participating in those activities. The results of the study indicated that students participating in swimming and boxing develop less motor ability as measured by the Larson Test than do students who participated in the other six activities.

TEST ITEMS

Baseball Throw for Distance. Larson identifies one factor of motor ability as gross body co-ordination, which he describes as, "The capacity of the individual to integrate movements of different kinds (different requirements for each phase of the activity) into one pattern." (6) The test Larson used to measure this factor of motor ability is the baseball throw for distance.

At the initial test phase of the experiment, on analysis of critical ratios between the means of the various activities only one significant difference was found. The difference, being between the weight-training and tumbling-gymnastics groups, was significant at the 5-per-cent level of confidence; weight training being the superior performance group.

Only one group, tumbling-gymnastics, made a gain that could be considered significant at a level of confidence as high as 5 per cent (Table 3). The lack of greater improvement in this particular test item may be due, in part, to the fact that the initial test came at the close of the baseball and softball season when many of the subjects tested had been participating in those activities; while the post-test came four and one-half months later. This is only assumption on the part of the experimenter and is not a tested fact. It is to be noted, however, that all groups, with the exception of swimming which showed a mean loss, did show a slight mean gain although too small to be considered significant.

Table 4 presents an analysis of mean gains for this test item. The improvement made by the tumbling-gymnastics group was significantly better than the conditioning, swimming, and tennis groups; while students participating in wrestling made significant gains over those students participating in swimming and conditioning. Since the critical ratio for significance of differences is a matter of degree; then the four activities—swimming, conditioning, tennis, and boxing—tend to develop to a lesser extent those gross body coordinations measured by the baseball throw for distance than do the other four activities studied.

Pull-ups. Chinning, or pull-ups, was the only item that was duplicated in both tests. As a test item in physical fitness, it is used as a measure of muscular endurance; or, "The capacity of the individual to continue successive exertions under conditions where a load is placed on the muscle groups being tested." (6) As a measure of motor ability, Larson used chinning as a test of a factor he described as dynamic strength which he defined as ". . . strength determined by the ability to lift the body weight and propel it upward. . ."(5).

In determining mean gain and significance of mean gain for each activity, it was found that two of the groups—wrestling and tumbling-gymnastics—made gains that were considered significant at the 1-per-cent level of confidence (Table 3). The mean gain of the weight-training group was significant at the 5-per-cent level of confidence. Two activities failed to show any increase in performance—swimming, the mean of which did not change, and boxing which showed an actual decrease in performance, although not a significant one.

The t-test for significance of differences in mean gains (Table 4) showed that the wrestling and tumbling-gymnastics groups made the most significant gains; while boxing and swimming seemed to develop this ability to a lesser degree than the other six activities studied.

Vertical Jump. In classifying tests in terms of the predominant factors which they measure, Cozens (3) classifies the vertical jump as a measure of jumping or leg strength. Larson (5) identifies it as a measure of a factor he terms motor explosiveness; in that the test item requires the use of the body weight in a force-time relationship.

TABLE 3

Mean Initial Scores, Gains, and Critical Ratios for Baseball Throw for Distance and Pull-Ups

Activity	N	Baseball Throw for Distance				Pull-Ups			
		Mean	S.D.	Gain	C.R.	Mean	S.D.	Gain	C.R.
Wrestling	58	177.94	24.82	7.08	1.56	6.86	2.45	1.55	3.88
Conditioning	50	179.30	27.23	2.40	0.45	6.82	2.44	0.70	1.43
Tennis	51	177.00	33.87	3.04	0.50	7.41	3.00	0.31	0.61
Volleyball	63	178.83	24.65	6.42	1.52	6.90	2.46	0.51	1.13
Swimming	52	176.13	23.29	—0.19	0.04	7.23	1.90	0.00	0.00
Weight training	47	184.10	20.07	5.67	1.34	7.51	2.25	0.96	2.09
Tumbling-gymnastics	60	174.92	20.52	8.41	2.33	7.70	2.27	1.30	2.17
Boxing	51	176.80	28.99	4.51	0.82	7.55	2.61	—0.28	0.67

TABLE 4

Critical Ratios of Differences in Mean Gains[1]

Activity	Wrestling	Conditioning	Tennis	Volleyball	Swimming	Weight Training	Tumbling-gymnastics	Boxing
Wrestling	——	2.17	1.60	0.31	2.90	0.66	0.66	1.03
Conditioning	2.83	——	0.23	1.68	0.94	1.33	2.57	0.77
Tennis	4.77	1.22	——	1.24	1.06	0.95	2.01	0.48
Volleyball	4.33	0.63	0.77	——	2.45	0.27	0.88	0.71
Swimming	5.17	2.00	0.97	1.70	——	2.14	3.26	1.56
Weight training	2.11	0.79	2.17	1.61	2.91	——	1.19	0.42
Tumbling-gymnastics	0.89	1.82	3.30	2.82	3.94	1.06	——	1.48
Boxing	7.04	3.06	2.11	3.04	0.88	4.13	5.27	——

[1] Figures above the dotted line in each column give ratios for the Physical Fitness Test; below, for the Larson Motor Ability Test.

In performance on the initial test, the groups were homogeneous in that no C.R. significant at the 5-per-cent level of confidence was found between the means of the various groups (Table 5). In studying mean gain and significance of mean gain for each activity, three groups—volleyball, tumbling-gymnastics, and boxing—made gains that were considered significant at the 1-per-cent level of confidence. Whereas all other activities showed a slight mean gain, no gains were of such proportion as to be considered significant.

A comparison of the mean gain of each group with the mean gain of each of the other groups (Table 6) indicated that the gain made by the tumbling-gymnastic group was significantly superior to the gain made by the other groups. The volleyball, conditioning, and tennis groups showed a statistically significant greater gain than did wrestling and weight-training groups. The data also indicated that the students who participated in weight training and wrestling least develop those abilities measured by this test item.

Bar-snap. The bar-snap, as reported by Larson correlated highly with dynamic strength; and Cozens describes it as a measure of body co-ordination, agility, and control.

In an analysis of the significance of differences between the means of the various groups on the initial test, two significant differences were found. The tennis group, on the initial test, was significantly superior to the conditioning group at the 1-per-cent level of confidence. The tennis group was likewise significantly superior to the wrestling group at the 5-per-cent level of confidence.

Each of the groups showed some measure of improvement between the means of the initial test and post-test (Table 5). The t-test for the significance of mean gain for each of the groups gave evidence that three groups—wrestling, conditioning, and tumbling-gymnastics—made gains that were of such proportion as to be considered significant at the 1-per-cent level of confidence. The mean gain made by the tennis group was significant at the 2-per-cent level of confidence; while the mean gains reported for the swimming and boxing groups were significant at the 5-per-cent level of confidence.

When the significance of mean gains was analylzed (Table 6), it was found that the gain made by the tumbling-gymnastics group was significantly superior to the gain made by each of the other groups except wrestling. The mean gain made by the wrestling group was significantly superior to the mean gains made by five groups; namely, tennis, volleyball, swimming, weight training, and boxing.

300-yard shuttle-run. This test item is described as a measure of muscular endurance, which is defined as,

> The capacity of the individual for long-continued contractions (submaximum) where a sufficient number of muscle groups are used with a sufficient duration and intensity to put a demand on the functions of circulation and respiration.

The means of the groups at the initial test phase ranged from m = 52.81 sec., weight training; to m = 53.88 sec., wrestling. The critical ratio of the difference in means of these two groups was significant at the 2-per-cent level of confidence.

TABLE 5

Mean Initial Scores, Gains, and Critical Ratios for Vertical Jump and Bar Snap

Activity	N	Vertical Jump				Bar Snap			
		Mean	S.D.	Gain	C.R.	Mean	S.D.	Gain	C.R.
Wrestling	58	20.04	2.17	0.60	1.36	54.28	4.77	5.63	4.73
Conditioning	50	19.36	4.20	1.11	1.61	53.64	2.09	3.52	6.13
Tennis	51	20.66	2.60	1.11	2.41	55.76	2.43	3.12	2.42
Volleyball	63	20.43	2.10	1.20	3.24	54.83	8.71	2.73	1.86
Swimming	52	20.17	1.92	0.98	2.51	53.94	8.07	3.14	2.00
Weight training	47	20.74	2.79	0.62	1.19	54.87	7.14	0.39	0.27
Tumbling-gymnastics	60	19.78	2.13	1.64	4.10	54.08	7.77	5.70	3.58
Boxing	51	20.16	2.05	1.04	2.89	54.33	6.31	2.98	2.01

TABLE 6

Critical Ratios of Differences in Mean Gains[1]

Activity	Wrestling	Conditioning	Tennis	Volleyball	Swimming	Weight Training	Tumbling-gymnastics	Boxing
Wrestling	—	2.32	2.13	2.73	1.58	0.08	4.33	1.83
Conditioning	1.80	—	0.00	0.45	0.59	2.23	2.41	0.32
Tennis	2.01	0.32	—	0.41	0.54	2.04	2.21	0.29
Volleyball	2.50	0.68	0.32	—	1.00	2.64	2.00	0.73
Swimming	2.13	0.32	0.02	0.35	—	1.50	2.75	0.25
Weight training	4.06	2.35	2.02	1.83	2.15	—	4.25	1.75
Tumbling-gymnastics	0.07	2.08	2.29	2.88	2.46	4.54	—	2.50
Boxing	2.19	0.45	0.11	0.21	0.13	1.96	2.50	—

[1] Figures above the dotted line in each column give ratios for the Physical Fitness Test; below, for the Larson Motor Ability Test.

An analysis of the mean gains (Table 7) between the initial test and post-test for each activity indicated that two groups—tennis and swimming—showed an actual performance loss, although too small to be considered more than chance. The other six activity groups all showed a performance gain between initial test and post-test means. The wrestling, conditioning, volleyball, and tumbling-gymnastics groups reported mean gains of sufficient size as to be considered significant at the 1-per-cent level of confidence. The mean gains of the other activities were not statistically significant. The 300-yard shuttle-run was practiced by those students taking the conditioning course.

In a comparison of the significance of differences between mean gains of the various groups (Table 8), the t-test indicated that the mean gain made by the conditioning group was significantly superior to the mean gains made by the swimming, weight-training, tennis, and boxing groups. Four other groups — wrestling, volleyball, boxing, and tumbling-gymnastics — showed mean gains that were significantly superior to the mean gains made by the swimming and weight-training groups. No other differences in mean gains were indicated as being significant at the 5-per-cent level of confidence.

Sit-ups. This test item was used as another measure of endurance, the purpose of which was to determine the capacity of the individual to perform successive exertions where a load is placed on certain muscle groups. In this exercise, the abdominal muscles are the predominant muscle groups tested.

There were no statistically significant differences between the means of the various activity groups on the initial test.

An analysis of the mean gain for each activity group between initial test and post-test performance (Table 7), five groups—wrestling, conditioning, tennis, volleyball, and tumbling-gymnastics—reported mean gains that were significant at the 1-per-cent level of confidence. The mean gain made by the boxing group was significant at the 5-per-cent level of confidence. No other significant gains were recorded.

A comparison of mean gains (Table 8), between each of the activity groups, indicated that the mean gain made by the conditioning group was proportionately greater to a significant degree than the mean gain reported by each of the other groups, with the exception of the tumbling-gymnastics group. The mean gain made by the tumbling-gymnastics group was significantly greater than the mean gain made by the tennis, volleyball, swimming, weight-training, and boxing groups. The mean gain reported for the wrestling group was significantly superior to the mean gains reported by the swimming, weight-training, and boxing groups. All activity groups, except swimming and boxing, reported mean gains that were significantly greater than the mean gain of the weight-training group. Of the eight activity groups studied, weight training seemed to develop to a lesser degree those phases of endurance necessary to perform repeated sit-up exercises.

TABLE 7

Mean Initial Scores, Gains, and Critical Ratios for 300-Yard Shuttle Run and Sit-Ups

Activity	N	300-Yard Shuttle Run				Sit-Ups			
		Mean	S.D.	Gain	C.R.	Mean	S.D.	Gain	C.R.
Wrestling	58	53.88	2.21	1.35	3.38	51.66	14.97	13.96	4.26
Conditioning	50	53.48	1.90	1.62	4.50	51.70	13.73	22.16	8.66
Tennis	51	53.67	2.25	—0.06	0.85	53.92	18.27	12.10	3.13
Volleyball	63	53.33	2.08	1.14	3.08	51.60	13.89	11.19	3.64
Swimming	52	53.48	2.12	—0.40	0.88	49.56	19.23	7.34	1.78
Weight training	47	52.81	2.03	0.19	0.46	51.15	13.85	5.28	1.76
Tumbling-gymnastics	60	53.81	2.10	1.36	3.40	53.50	15.58	18.75	5.14
Boxing	51	53.61	2.66	0.85	1.81	53.67	15.77	7.37	2.30

TABLE 8

Critical Ratios of Differences in Mean Gains[1]

Activity	Wrestling	Conditioning	Tennis	Volleyball	Swimming	Weight Training	Tumbling-gymnastics	Boxing
Wrestling	.....	0.73	3.62	0.66	4.86	3.22	0.03	1.43
Conditioning	2.48	.....	4.31	1.50	5.61	4.00	0.74	2.20
Tennis	0.72	3.27	.....	3.64	0.92	0.68	3.94	2.53
Volleyball	0.99	3.38	0.34	.....	5.13	3.17	0.79	1.04
Swimming	2.26	4.40	1.78	1.34	.....	1.69	5.33	3.79
Weight training	3.19	5.29	2.80	2.23	0.74	.....	3.55	2.00
Tumbling-gymnastics	1.53	0.96	2.29	2.45	3.55	4.46	.....	1.59
Boxing	2.14	4.23	1.67	1.26	0.01	0.71	3.41	.....

[1] Figures above the dotted line in each column give ratios for the Physical Fitness Test; below, for the Larson Motor Ability Test.

Summary

The purpose of this study was to investigate the relative effect of several physical education activities on motor ability and physical fitness. The activities used in the study were: boxing, conditioning, swimming, tennis, tumbling-gymnastics, volleyball, weight training, and wrestling. The tests used as the criterion measure to determine motor ability and physical fitness were the Larson Test of Motor Ability and the Physical Fitness Test, which was a slight modification of the Army Air Forces PFR test.

The study was conducted during the fall semester of the academic year 1950-51 at the Agricultural and Mechanical College of Texas, the subjects for the experiment being selected from 1,031 male first semester Freshmen at that institution. All incoming Freshmen were tested by the Physical Fitness Test; and students were selected for the experiment on the basis of homogeneity of two factors, scores of the Physical Fitness Test and chronological age. The selected group was then divided into eight groups, keeping the groups as equivalent in physical fitness and chronological age as possible.

A third factor was used in the selection of the swimming group, namely, all students assigned to this group had failed rather rigid swimming tests. All groups were then tested for motor ability by the Larson Test of Motor Ability.

Each group was then assigned to one of the designated activities in which they participated for one semester. At the end of the semester all groups were again tested by the same two tests. To reduce outside variables, in physical education activities, all students were eliminated from the study who participated in intramural activities more than three hours during the semester.

The data consisted of records of performance on the individual test items used and total scores for the two tests at the initial test and post-test stages. Statistical treatment of the data was made to calculate means and standard deviations for each test item and total test scores as recorded at both testing periods. By use of the critical ratio, the significance of difference in mean gains was studied to determine if there was a significant difference in the relative effect of the various activities on physical fitness and motor ability.

Conclusions

The results of these data indicated that improved physical fitness rating was equally well attained by those students participating in conditioning and by those students who participated in tumbling-gymnastics; whereas swimming, tennis, and boxing are the activities least apt to increase the student's score on the Physical Fitness Test. It was further found from the study that the groups participating in wrestling and tumbling-gymnastics made the most significant gains on the Larson Test of Motor Ability. Again, the groups indicating the least improvement were tennis, swimming, and boxing. Tumbling-gymnastics ranked equal to conditioning in the development of physical fitness and equal to wrestling in the development of motor ability; while wrestling and conditioning each ranked high on only one of the two tests.

Further, the tumbling-gymnastics group was the only group that evidenced a significant gain between the initial test and post-test means for each of the individual test items of both measures used. The other seven groups failed to record a significant gain between the pre-test and post-test means in at least two or more of the test items.

The results would seem to indicate that the course using the combined activities of tumbling and gymnastics best developed all phases of physical fitness and motor ability tested. The data also indicated that the activities which tend to least develop these abilities in students are tennis, swimming, and boxing.

The investigation does not mean to imply that the development of specific sports skills in any of the recreational areas is secondary to the development of physical fitness or motor ability as measured by the tests used. It cannot be said, for example, that the ability to score high on the measures used in this study is more important than learning the skills necessary to participate successfully in activities that are satisfying to the individual. Each has its place in the physical education program.

REFERENCES

1. Army Air Forces Regulation No. 50-10. *Training-Physical Fitness Test.* Washington, D. C., Headquarters, Army Air Forces, April 28, 1943.
2. Berrafato, Peter R. *The Effect of Various Physical Education Service Courses on the All-Round Muscular Endurance of University Students.* Unpublished Master's thesis, University of Illinois, 1949.
3. Cozens, Frederick W. *The Measurement of General Athletic Ability in College Men.* Eugene, Oregon: University Press, 1929, p. 132.
4. Fordham, Sheldon L. *The Effect of Four Selected Physical Education Activities on Muscular Endurance Test Scores.* Unpublished Master's thesis, University of Illinois, 1949.
5. Larson, Leonard A., A Factor Analysis of Motor Ability Variables and Tests, with Tests for College Men. *Research Quarterly,* **12**: 499-517, 1941.
6. ———, Some Findings Resulting for Army Air Forces Physical Training Program. *Research Quarterly,* **17**: 147, 1946.
7. Wilbur, E. A., A Comparative Study of Physical Fitness Indices as Measured by Two Programs of Physical Education: The Sports Method and the Apparatus Method. *Research Quarterly,* **14**: 316-332, 1943.

(Submitted 8/13/54)

Summary

The purpose of this study was to investigate the relative effect of several physical education activities on motor ability and physical fitness. The activities used in the study were: boxing, conditioning, swimming, tennis, tumbling-gymnastics, volleyball, weight training, and wrestling. The tests used as the criterion measure to determine motor ability and physical fitness were the Larson Test of Motor Ability and the Physical Fitness Test, which was a slight modification of the Army Air Forces PFR test.

The study was conducted during the fall semester of the academic year 1950-51 at the Agricultural and Mechanical College of Texas, the subjects for the experiment being selected from 1,031 male first semester Freshmen at that institution. All incoming Freshmen were tested by the Physical Fitness Test; and students were selected for the experiment on the basis of homogeneity of two factors, scores of the Physical Fitness Test and chronological age. The selected group was then divided into eight groups, keeping the groups as equivalent in physical fitness and chronological age as possible.

A third factor was used in the selection of the swimming group, namely, all students assigned to this group had failed rather rigid swimming tests. All groups were then tested for motor ability by the Larson Test of Motor Ability.

Each group was then assigned to one of the designated activities in which they participated for one semester. At the end of the semester all groups were again tested by the same two tests. To reduce outside variables, in physical education activities, all students were eliminated from the study who participated in intramural activities more than three hours during the semester.

The data consisted of records of performance on the individual test items used and total scores for the two tests at the initial test and post-test stages. Statistical treatment of the data was made to calculate means and standard deviations for each test item and total test scores as recorded at both testing periods. By use of the critical ratio, the significance of difference in mean gains was studied to determine if there was a significant difference in the relative effect of the various activities on physical fitness and motor ability.

Conclusions

The results of these data indicated that improved physical fitness rating was equally well attained by those students participating in conditioning and by those students who participated in tumbling-gymnastics; whereas swimming, tennis, and boxing are the activities least apt to increase the student's score on the Physical Fitness Test. It was further found from the study that the groups participating in wrestling and tumbling-gymnastics made the most significant gains on the Larson Test of Motor Ability. Again, the groups indicating the least improvement were tennis, swimming, and boxing. Tumbling-gymnastics ranked equal to conditioning in the development of physical fitness and equal to wrestling in the development of motor ability; while wrestling and conditioning each ranked high on only one of the two tests.

Further, the tumbling-gymnastics group was the only group that evidenced a significant gain between the initial test and post-test means for each of the individual test items of both measures used. The other seven groups failed to record a significant gain between the pre-test and post-test means in at least two or more of the test items.

The results would seem to indicate that the course using the combined activities of tumbling and gymnastics best developed all phases of physical fitness and motor ability tested. The data also indicated that the activities which tend to least develop these abilities in students are tennis, swimming, and boxing.

The investigation does not mean to imply that the development of specific sports skills in any of the recreational areas is secondary to the development of physical fitness or motor ability as measured by the tests used. It cannot be said, for example, that the ability to score high on the measures used in this study is more important than learning the skills necessary to participate successfully in activities that are satisfying to the individual. Each has its place in the physical education program.

REFERENCES

1. Army Air Forces Regulation No. 50-10. *Training-Physical Fitness Test.* Washington, D. C., Headquarters, Army Air Forces, April 28, 1943.
2. Berrafato, Peter R. *The Effect of Various Physical Education Service Courses on the All-Round Muscular Endurance of University Students.* Unpublished Master's thesis, University of Illinois, 1949.
3. Cozens, Frederick W. *The Measurement of General Athletic Ability in College Men.* Eugene, Oregon: University Press, 1929, p. 132.
4. Fordham, Sheldon L. *The Effect of Four Selected Physical Education Activities on Muscular Endurance Test Scores.* Unpublished Master's thesis, University of Illinois, 1949.
5. Larson, Leonard A., A Factor Analysis of Motor Ability Variables and Tests, with Tests for College Men. *Research Quarterly,* **12**: 499-517, 1941.
6. ———, Some Findings Resulting for Army Air Forces Physical Training Program. *Research Quarterly,* **17**: 147, 1946.
7. Wilbur, E. A., A Comparative Study of Physical Fitness Indices as Measured by Two Programs of Physical Education: The Sports Method and the Apparatus Method. *Research Quarterly,* **14**: 316-332, 1943.

(Submitted 8/13/54)

Vitamin B_{12} : A Review

HENRY J. MONTOYE
Michigan State College
East Lansing, Michigan

Abstract

Recently, there has been considerable interest among coaches in the effects of vitamin B_{12} on athletic performance. This review article summarizes the research reported on the source and requirements of vitamin B_{12} in the diet, the role of this vitamin in red blood cell formation and metabolism, the effects of vitamin B_{12} on growth and work performance, and the relative efficiency of oral and parenteral feeding of this nutrient.

DESPITE THE FACT that evidence for securing improvement in work performance by dietary supplementation in normal, reasonably well-fed human beings is not encouraging (23, 42), interest in this approach does not seem to wane. One substance employed very recently for this purpose is Vitamin B_{12}. Two amateur athletic teams and one professional boxer are known to the author to use this nutrient regularly. Also, the vitamin has been recommended as a diet supplement for underweight athletes (40). Yet the scientific basis for this practice has not been established. Because it appears important for physical educators and health educators to have at their disposal a summary of pertinent literature on Vitamin B_{12}, this review is presented.

With the observation of Minot and Murphy in 1926 (44) that pernicious anemia can be aided by liver feeding, the quest was begun for the specific factor or factors involved. This work finally resulted in the isolation of Vitamin B_{12}, a water-soluble vitamin, in England and the United States at about the same time (44).

Source and Requirements

Viamin B_{12} is one of the most recently isolated dietary requirements and also one of the most potent therapeutic nutrients known. It is estimated to have 10,000 times the potency of refined liver extract in treating pernicious anemia (20). The chemical formula for this substance or its various forms (Vitamin B_{12a}, B_{12b}, B_{12c}) need not concern us here except possibly we might point out that the molecule contains carbon, hydrogen, nitrogen and oxygen plus a heavy metal, cobalt, and phosphorus. Because of the cobalt, biological isotope studies are practical. Natural B_{12} is not found in vegetable tissue. It is synthesized by micro-organisms exclusively (43), probably in the rumen of animals (50). Beef liver and kidney are rich sources of the vitamin and some is contained in other meats, egg yolk, cheese, and casein (43, 48).

The human requirement is probably only about one to two micrograms per day but healthy subjects apparently are not saturated with the vitamin since as much as 55 micrograms have been retained after a massive dose (4). Deficiencies are believed to be rare among healthy men and women (24). This

would seem to be especially true among athletes since they are frequently heavy meat eaters.

Despite the potency of this vitamin, it is relatively non-toxic. As much as 1,000 micrograms by injection or 10,000 micrograms given orally in one dose have not resulted in toxicity (4).

Vitamin B_{12} and Red Blood Cell Formation

Hemoglobin contained in the red blood cell is important in athletics, particularly because of its role in the transfer of oxygen from lungs to tissues and carbon dioxide from tissues to lungs. Vitamin B_{12} is capable of raising the hemoglobin level of the blood but these effects have been demonstrated only in cases of pernicious anemia (6, 18, 20, 30, 48), macrocytic anemia (20), sprue[1] (6), and among malnourished children (46). In normal subjects the vitamin had no effect on hemoglobin, red blood cell count, or eosinophil level (33). In experimentally-anemic B_{12}-deficient animals, the vitamin elicited a blood response (13, 17, 34, 35, 51) but in normal animals the hemoglobin was not affected (25). It is generally believed that vitamin B_{12} is necessary in the maturation of red blood cells from reticulocytes, the precursor of erythrocytes (2, 48, 50). One writer contends that B_{12} is used directly in red blood cell formation and not as a catalyst (36).

In addition to those diseases in which the beneficial effects of B_{12} have been established, Wetzel (54) claims to have been successful in treating one case of asthma with vitamin B_{12}, but this was not confirmed in a later study (21). Temporary relief of trigeminal neuralgia with B_{12} has been reported (49). Nerve regeneration and hence better co-ordination have also been reported as a result of B_{12} supplementation in pernicious anemia patients (15).

Vitamin B_{12} and Metabolism

Amino acids derived from the breakdown of consumed protein are the building blocks of body tissue. One of these essential amino acids is methionine. It is now well established that vitamin B_{12} has a decisive role in the bio-synthesis of methyl groups (3, 7, 9, 10, 12, 18, 22, 26, 37, 41, 47). Probably the B_{12} is needed as a co-enzyme in the manufacture labile methyl groups. This would explain the increase in B_{12} requirement when a high protein diet is taken. In one study, Vitamin B_{12} was reported to be necessary in the metabolism of glycine; another, amino acid (29). Absence of Vitamin B_{12} resulted in early death among animals on a high glycine diet. Another interesting application of B_{12}, related to its tissue-building function is that of decreasing wound-healing time (14).

In addition to its established importance in protein metabolism, B_{12} is believed to be involved in carbohydrate and possibly fat metabolism (4). It apparently increases carbohydrate utilization, since rats which were deficient

[1]A disease in which the blood picture is very similar to that which occurs in pernicious anemia.

in this nutrient had smaller carbohydrate reserves and intravenous glucose was used poorly (28). In another animal study, high blood sugar in the glucose tolerance test was also noted (27). An interesting experiment reported by Edwards and Carter (11) utilized rats fed a protein-poor diet which resulted in a negative nitrogen balance and high fat deposits in the liver. Feeding 50 micrograms of methionine brought the nitrogen to equilibrium but the animals continued to have fatty livers. When B_{12} was added to the methionine, the liver became normal and a 50 per cent increase in production of hemoglobin was noted. It is now known that methyl synthesis cures the fatty livers.

Vitamin B_{12} and Growth

Probably because of its role in protein metabolism, Vitamin B_{12} has been demonstrated to be essential to normal growth among various animals (3, 19, 41, 45). The usual experimental procedure in these studies has been to feed animals a diet deficient in B_{12} and then add the vitamin to part of the group, keeping the others as controls on the B_{12}- deficient diet. An interesting observation in one study of chicks was the myoatrophy of the legs, which was very marked in B_{12} deficiency (38).

Studies on human beings have yielded conflicting results. In normal adults no significant effects on weight were noted with B_{12} supplementation (1). B_{12} also had no effect on the growth of premature infants (8, 39, 44, 46). In a well-controlled study of children selected as having "simple growth failure," oral feeding of Vitamin B_{12} had a marked stimulatory effect on growth (53). Some studies have confirmed this (4, 54) but several did not (46, 52). However, evidence in general points to improvement in growth with B_{12} in children, provided those selected might be expected to have a deficiency in this nutrient, there is *ad libitum* access to food during treatment and there is adequate nutrient and caloric distribution in the diets offered (16).

Vitamin B_{12} and Work Performance

Wetzel (53) observed no effect on leg, back, or grip strength among the malnourished children who benefited in growth from B_{12} feeding. In normal young men, Montoye *et al.* (32) were able to show no effects of B_{12} supplementation on grip strength, pulse rate recovery after moderate exercise or maximum performance on a bicycle ergometer. In a later study (31) with young institutionalized boys, no effects of B_{12} were observed on the step-test score or half-mile-run time. Both the control and experimental groups lost weight in this study, but, interestingly enough, the thinner boys (B-1 Channel, Wetzel Grid) in the control group lost twice as much weight and exhibited four times as much channel shift as the thin boys on Vitamin B_{12}. The number of cases, however, was too small to consider this anything but an indication of a need for repeating the study.

The lack of effect on performance is not surprising in the light of what is known about Vitamin B_{12} and the lack or rarity of deficiencies of this vita-

min among healthy subjects. There does not seem to be sufficient evidence to recommend Vitamin B_{12} supplementation as a general practice among athletic teams.

Oral Versus Parenteral Feeding

There is little doubt that the injection method of supplying Vitamin B_{12} is more efficient than oral feeding (4, 5, 6, 50, 55). The injection dose for pernicious anemia patients is one gamma per day by injection, but orally it is 50 to 75 gamma per day. However, injections are not practical when medical supervision is not available, as is frequently the case in athletics, industry, and military operations. Furthermore, there is ample evidence that oral administration results in B_{12} activity in the urine (4, 55) and higher plasma values (55). Evidence that oral doses are effective among deficient animals and human beings is available in the many controlled studies quoted earlier in which growth was stimulated. Blood responses have been secured with oral administration in many independent studies with cases of anemia. The inclusion of the intrinsic factor reduced the required oral doses in anemia cases (2, 6), but that probably can be explained by a deficiency of this factor in pernicious anemia. It is commonly believed that Vitamin B_{12} is used directly from food and the intrinsic factor of gastric juice aids in the absorption and not the synthesis of Vitamin B_{12} (48). It is doubtful that the administration of gastric juice would increase absorption in normal subjects.

REFERENCES

1. Bayne, G. M., and W. P. Boger, Vitamin B_{12} supplementation of the diet of healthy adults: a study of weight changes employing a latin square design, *J. Clinical Nutrition*, **1**: 424-9, 1953.
2. Callender, S. T., and L. G. Lajtha, Nature of Castles hemopoietic factor, *Blood*, **6**: 1234-9, 1951.
3. Charkey, L. W., *et al.*, Vitamin B_{12} in amino acid metabolism, *Proc. Soc. Exper. Biol. Med.*, **73**: 21-24, 1950.
4. Chow, B. F., The role of Vitamin B_{12} in metabolism, *Southern Medical J.* **45**: 604-612, 1952.
5. ———, *et al.*, The appearance of vitamin B_{12} activity in urine after oral and intramuscular administration to man, *Bulletin Johns-Hopkins Hospital*, **87**: 156-163, 1950.
6. Conley, C. L., *et al.*, Observations on the Absorption, Utilization, and Excretion of Vitamin B_{12}, *J. Laboratory and Clinical Medicine*, **38**: 84-94, 1951.
7. Davis, B. D., and E. S. Mingioli, Mutants of Escherichia Coli requiring methionine or vitamin B_{12}, *J. Bacteriology*, **60**: 17-28, 1950.
8. Downing, D. F., Failure of vitamin B_{12} to promote growth of premature infants, *Science*, **112**: 181, 1950.
9. Dubnoff, J. W., The role of B_{12} in methionine synthesis in E. Coli, *Archives of Biochemistry and Biophysics*, **37**: 37-45, 1952.
10. ———, Vitamin B_{12} in methionine formation, *Federation Proc.* **10**: 178, 1951.
11. Edwards, C. H., and R. T. Carter, Interrelationships between vitamin B_{12} and methionine in metabolism, *Federation Proc.* **12**: 413, 1953.
12. Emerson, G., and K. Folkers, Water-soluble vitamins, *Annual Review of Biochemistry*, **20**: 559-598, 1951.

13. FATTERPAKER, P., V. MARFATIA, and A. SREENIVASAN, Influence of folic acid and vitamin B_{12} on formation of creatine *in Vitro* and *in Vivo*, *Nature*, **167**: 1067-8, 1951.
14. FINDAY, C. W., Effect of vitamin B_{12} on wound healing, *Proc. Soc. Exper. Biol. Med.*, **82**: 492, 1953.
15. HALL, B. E., F. H. KRUSEN, and H. W. WOLTMAN, Vitamin B_{12} and coordination exercises for combined degeneration of the spinal cord in pernicious anemia, *J. of the American Medical Association*, **141**: 257-260, 1949.
16. HARTE, R. A., and B. F. CHOW, Vitamin B_{12} conference, *Science*, **118**: 582-3, 1953.
17. JOHNSON, B. C., *et al.*, The interrelationship of vitamin B_{12} and folic acid in the baby pig," *J. Laboratory and Clinical Medicine*, **36**: 537-46, 1950.
18. JUKES, T. H., and E. L. R. STOKSTAD, The role of vitamin B_{12} in metabolic processes, *Vitamins and Hormones*, **9**: 1-26, 1951.
19. ———, E. L. R. STOKSTAD, and H. P. BROQUIST, Editorial, *Archives of Biochemistry*, **25**: 453-5, 1950.
20. KAHN, S. S., and F. J. STARE, Nutrition in medicine, *Annual Review of Medicine*, **1**: 127-152, 1950.
21. KAUFMAN, R. E., Effect of vitamin B_{12} in asthma, *Annals of Allergy*, **9**: 517-8, 1951.
22. KAYE, R., *et al.*, The minimal nitrogen requirement of infants and the effect of vitamin B_{12} upon it, *American J. Diseases of Children*, **86**: 326-7, 1953.
23. KEYS, ANCEL, Physical Performance in relation to diet, *Federation Proc.* **2**: 164-187, 1943.
24. LANG, C., *et al.*, Retention of crystalline vitamin B_{12} by healthy male individuals following intramuscular injection, *J. Nutrition*, **46**: 215-221, 1952.
25. LEVEY, S., and J. M. ORTEN, Vitamin B_{12} and the production of polycythemia by cobalt, *J. Nutrition*, **45**: 487-492, 1951.
26. LIENER, I. E., and M. O. SCHULTZE, The excretion of n'-methylnicotinamide and creatine as influenced by vitamin B_{12} and various methyl donors, *J. of Nutrition*, **46**: 223-237, 1952.
27. LING, C. T., and B. F. CHOW, Effect of vitamin B_{12} on carbohydrate and fat metabolism, *J. Biological Chemistry*, **206**: 797-805, 1954.
28. ———, Metabolic role of vitamin B_{12}, *Federation Proc.* **11**: 249, 1952.
29. MENCH, H., and G. F. COMBS, Action of vitamin B_{12} in counteracting glycine toxicity in the chick, *Proc. Soc. Exper. Biol. Med.*, **75**: 139-142, 1950.
30. MEYER, L. M., Oral treatment of pernicious anemia with vitamin B_{12}, *American J. of the Medical Sciences*, **220**: 604-609, 1950.
31. MONTOYE, HENRY J., *et al.*, The effect of vitamin B_{12} on endurance, Paper read before the Research Section, 58th National Convention, AAHPER, New York, April 21, 1954.
32. ———, Effect of vitamin B_{12} on work capacity (in press).
33. MUEHRCKE, R. C., and R. M. KARK, Negative effects of vitamin B_{12} on blood of healthy individuals with special reference to eosinophils, *Proc. Soc. Exper. Biol. and Med.*, **77**: 144-5, 1951.
34. NEWMAN, A. L., B. C. JOHNSON, and T. B. THIERSCH, Crystalline vitamin B_{12} in the nutrition of the baby pig, *J. Nutrition*, **40**: 403-414, 1950.
35. NICHOL, C. A., A. E. HARPE and C. A. ELVEHJEM, Effect of folic acid, liver extract and vitamin B_{12} on hemoglobin regeneration in chicks, *Proc. Soc. Exp. Biol. Med.*, **71**: 34-7, 1949.
36. OCKRENT, C., Relation between vitamin B_{12} and the red blood cells, *Nature*, **165** 280-281, 1950.
37. OGINSKY, E. L., Editorial, *Archives of Biochemistry*, **26**: 327-9, 1950.
38. OLCESE, O., *et al.*, Congenital anomalies in the chick due to vitamin B_{12}, *J. Nutrition*, **41**: 423-431, 1950.

39. RASCOFF, H. A. DUNEWITZ and R. NORTON, The weight progress of premature infants given supplementary feedings of vitamin B_{12}, *J. Pediatrics*, **39**: 61-4, 1951.
40. *Research Guide*, Nov. 1952, p. 2, published by Sports College, Toronto, Canada.
41. SCHWEIGERT, B. S., Significance of vitamin B_{12} and related factors, *J. American Dietetic Association*, **26**: 782-7, 1950.
42. SIMONSON, E., Influence of nutrition on work performance, *Nutrition Fronts in Public Health*, pp. 72-117, Nutrition Symposium Series No. 3, National Vitamin Foundation, Inc., 1951.
43. SMITH, E. L., The discovery and identification of vitamin B_{12}, *British J. Nutrition*, **16**: 295-9, 1952.
44. ———, Vitamin B_{12}: Part I, *Nutritional Abstracts and Reviews*, **20**: 795-809, 1951.
45. ———, The vitamin B_{12} group of factors, *British Medical J.* **1**: 151-2, 1951.
46. SPIES, T. D., *et al.*, The effect of supplements of vitamin B_{12} on the growth of children with nutritive failure, *Internationale Zeitschrift für Vitaminforschung* **23**: 414-29, 1952.
47. STEKOL, J. A., and K. WEISS, Vitamin B_{12} and growth of rats on diets free of methionine and choline, *J. Biological Chemistry*, **186**: 343-350, 1950.
48. STRAUSS, M. B., Vitamin B_{12} and pernicious anemia, *New England J. Medicine*, **243**: 187-194, 222-9, 1950.
49. SURTEES, S. J., and R. R. HUGHES, Treatment of trigeminal neuralgia with vitamin B_{12}, *Lancet*, **1**: 439-441, 1954.
50. UNGLEY, C. C., Vitamin B_{12}. Part 2: a review of the clinical aspects, *Nutrition Abstracts and Reviews*, **21**: 1-26, 1951.
51. VIJAYARAGHAVAN, P. K., and M. S. DUNN, Effect of vitamins B_{12a} and B_{12b} on the red blood cell counts in experimental anemia in mice, *Archives of Biochemistry*, **31**: 248-250, 1951.
52. Vitamin B_{12} and growth, Editorial, *British Medical J.*, **1**: 1401, 1952.
53. WETZEL, N. C., *et al.*, Growth failure in school children: further studies of vitamin B_{12} dietary supplements, *J. Clinical Nutrition*, **1**: 17-31, 1952.
54. ———, *et al.*, Growth failure in school children as associated with vitamin B_{12} deficiency-response to Oral therapy, *Science*, **110**: 651-3, 1949.
55. YAMAMOTO, R., *et al.*, Further studies on B_{12} absorption following oral and parenteral administration, *J. Nutrition*, **45**: 507-519, 1951.

(Submitted 11/1/54)

Analysis of Results from the Kraus-Weber Test of Minimum Muscular Fitness in Children

MARJORIE PHILLIPS, CAROLYN BOOKWALTER, CHARLOTTE DENMAN, JANET McAULEY, HILDA SHERWIN, DEAN SUMMERS, AND HELEN YEAKEL

Indiana University
Bloomington, Indiana

Abstract

The Kraus-Weber Test of muscular fitness was administered to a population of 1456 elementary school children. Results indicate:

1. The Test and Test items are highly reliable.
2. There is no relationship between grip strength and success on the Test.
3. Girls are substantially superior to boys at all levels in passing the Test, and in flexibility.
4. For both sexes, there is a decided decrease in strength item failure as age increases. By the time the children are 11 years of age the percentage of failure on any item is less than 8 per cent, and agrees very closely with the results found for European children.

IN THE MAY 1954 issue of the *Research Quarterly* and the December 1953 issue of the *Journal of Health, Physical Education, and Recreation*, Dr. Hans Kraus and Mrs. Ruth P. Hirschland reported the results of their research on the muscular fitness of children. The information revealed in these reports should have aroused the concern of every person interested in the health and fitness of the young people in our Country. The apparent deficiencies of American children as compared to the children of other countries present an alarming picture.

It was the purpose of this investigation to measure a city population of Indiana children in an attempt to determine how their muscular fitness test scores compared with the scores of the children measured by Dr. Kraus and Mrs. Hirschland.

Other purposes of the investigation were to determine the reliability of the Kraus-Weber Test, to compare the Test results to results from a test of dynamometric strength, and to analyze performances on each of the Test items by age and by sex.

The city chosen for the investigation can reasonably be considered a typical midwestern city for its population size of 26,000. It has four public elementary schools and all of the physically normal population of these schools, grades one through six, who were present in school on the days the investigators visited, were given the muscular fitness tests. Rural schools within the

City School District, religious schools and laboratory schools were not included in the study.

Procedures

Seven faculty members of the Indiana University School of Health, Physical Education, and Recreation composed the team which administered all tests. One member of the team had been trained and certified by Dr. Kraus in the administration of the Kraus-Weber Test. Ths team member was responsible for supervising and instructing other members of the team in proper testing procedures. A preliminary training period was held for this purpose.

The seven testors worked as a team througout the entire testing period. In some schools a special testing room was assigned with seven tables in the room and the subjects came there to be tested. In other schools, the team moved from room to room and used the double station desks available, for purposes of giving the tests.

In addition to the original administration, repeats of the test were given in order to determine test reliability. All repeats were given within five to fifteen minutes of the original test. Repeats were so distributed that schools, ages, and sexes were reasonably represented. The repeat of the test was always done independently by some other testor than the one administering the original test. A new scoring card with complete information was used for each repeated test, and these cards were matched up with the original ones during the subsequent analysis of the results.

A manuometer[1] was used to determine the grip strength of a sample of children. Grip strengths were secured from the fourth, fifth, and sixth-grade children of one of the schools. The younger children were excluded, since the manuometer is rather difficult for them to handle.

Test Description

The Kraus-Weber Test is composed of six items, with failure on any item interpreted as a Test failure. The six items are described briefly as follows:[2]

1. *Flexibility*—ability to bend over slowly and touch the finger tips to the floor without bending the knees. 2. *Abdominals plus psoas*—supine position, legs straight and hands behind the neck. Ability to sit up with support being given at the ankles. 3. *Abdominals minus psoas*—the same as abdominals plus psoas, only with the knees bent, feet on table. 4. *Psoas*—supine position,

[1]The standard grip dynamometer was used, which provides no adjustment for variations in hand size. Purchasable from the Narragansett Gymnasium Equipment Company, Centralia, Missouri.

[2]For a complete description of the Test, see Kraus, Hans, M.D., and Hirschland, Ruth P. Minimum muscular fitness tests in school children. *Research Quarterly*, **25**: 178-188, May 1954.

The description of the muscle actions involved in the performance of the various test items represent the opinions of Kraus and Weber, and it was not the intent of this research to evaluate their analysis.

hands behind neck, legs straight. Ability to raise the legs with feet approximately ten inches off the table and hold for ten seconds. 5. *Upper back*—prone position, hands behind the neck, supporting pillow under the abdomen. Ability to raise the trunk and hold for ten seconds, with operator holding feet down. 6. *Lower back*—prone position, knees straight, arms on table, supporting pillow under abdomen. Ability to raise the legs and hold for ten seconds, with operator holding upper back down.

The last five items are considered by Kraus and Weber to be strength tests of the muscle groups concerned. One trial is given on each item of the Test, and the subject is passed or failed on this basis.

Analysis and Discussion of Data

The tests were administered to a total of 1,473 children, ages 6 through 15. Since only 17 of the children were over 12 years of age, these 17 were eliminated in reporting the results. Thus, the report deals with 1,456 children, ages 6 through 12.

RELIABILITY OF THE TESTS

The tests were repeated with 215 children, and analysis of the results reveals the tests to be highly reliable. Since the Test itself is scored as either passed or failed, as is each item of the Test, the reliability analysis necessitated the use of a technique adapted to a double dichotomy.

On two of the items, upper back and lower back, there was complete agreement between the first and second adminstrations. On other items and the Test itself, the degree of agreement was so high that the tetrachoric coefficient and the Phi coefficient both estimated the relationship to be slightly over one. Obviously, neither of these techniques was suitable, since the degree of agreement, while very high, was not perfect. A third technique, which may be used in estimating relationship from a two by two table, is Yule's Q (4, *p. 210*). This technique yielded the results given in Table 1.

TABLE 1

Reliability Coefficients Determined by Yule's Q for the Kraus-Weber Test and Four of the Items of the Test

Kraus-Weber Test	Flexibility	Abdominals plus psoas	Abdominals minus psoas	Psoas
.954	.954	.986	.974	.990

Since all of these coefficients are above .950, it can be seen that the tests are remarkably reliable, particularly in view of the fact that seven different people were involved in making the judgments.

GRIP STRENGTH RESULTS

There is substantial evidence in the literature (4, *p. 89*; 5, *p. 503*; 7, p. *65*; 8, *p. 768*) to the effect that grip strength shows a considerable relationship to the general muscular strength of the body, as measured by the dynamometer. Correlations between grip strength and general strength factors range as high as from .650 to .763.

The grip strength of the preferred hand of a sample of 126 children, 61 boys and 65 girls, was measured by the manuometer. This sample included all of the girls and boys in the fourth, fifth, and sixth grades of one of the schools.

The mean grip strength of the 30 boys who passed the Kraus-Weber Test was compared to the mean grip strength of the 31 boys who failed the Kraus-Weber Test. The mean grip strength of the "passes" was found to be 45.6 pounds and of the "failures" 46.1 pounds. The difference between the means was evaluated by the t test, which indicated that the difference lacked statistical significance ($t = .23$). The same comparison was made for the girls, in which the 35 "passes" had a mean grip strength of 39.2 pounds and the 30 "failures" a mean grip strength of 42.7 pounds. The difference between these two means also lacked statistical significance ($t = 1.30$).

The grip strength of all of the boys was compared to that of all of the girls. The mean grip strength for boys was found to be 45.9 pounds and that for girls 40.8 pounds. The difference between these means was found to be significant at better than the 1-per-cent level of confidence ($t = 2.83$).

The mean ages for the four groups—pass and fail boys, and pass and fail girls—were very similar, not varying more than one-tenth of a year.

These results would appear to indicate that the Kraus-Weber Test is not related to general muscular strength insofar as the grip strength test can be considered to be measuring general muscular strength. These results are not surprising, however, since the majority of failures in the sample were deficient on the flexibility phase of the Kraus-Weber Test. The small number of muscle weakness failures made a comparison of just those failures hardly worthwhile.

COMPARISON OF INDIANA RESULTS WITH KRAUS RESULTS

Table 2 presents a comparison between the results obtained by Kraus and Hirschland in the population they measured and the population measured by the Indiana investigators. For purposes of convenience in drawing comparisons, the populations will be referred to as the Kraus and Indiana Groups.

TABLE 2

Comparison of Failures in Total Populations, Kraus and Indiana Groups

Group	Per Cent Incidence of Failures	Per Cent Test Failures	Per Cent Flexibility Failures	Per Cent Incidence Weakness Failures
Kraus	80.0	57.9	44.3	35.7
Indiana	53.9	45.1	36.3	17.7

A glance at the figures reveals immediately that the Indiana group is somewhat superior to the Kraus group in all comparisons. The most striking differences occur in incidence comparisons.

In arriving at incidence figures for the Test, the number of flexibility failures, abdominal failures, psoas failures and back failures are summed. The per cent that this sum is of the total Tests administered yields the incidence

percentages. A child is counted as an abdominal failure if he fails on either or both of the abdominal tests and a back failure if he fails on either or both of the back tests. Thus, a child could be counted four times if he were to fail on each of the four areas of the Test. Incidence of weakness failures is found in the same manner, except the flexibility item is eliminated from the count; thus, a child could be counted three times if he failed in each of the weakness areas.

The figures in Table 2 reveal that the Kraus group contained larger percentages of multiple failures on the Test than did the Indiana group, and that multiple failures in the strength test areas are apparently contributing more to the wide difference in the incidence figures. The difference in per cent of flexibility failures for the two groups is only 8.0 while the difference in per cent of incidence of weakness failures is 18.0. It is not possible to compare the two groups in terms of number of children with weakness failures, since those figures are not included in the Kraus report.

The comparison of failures at different age levels is given in Table 3.

TABLE 3

Comparison of Failures at Different Age Levels in Kraus and Indiana Groups*

Type Failure	Group	Age 6	7	8	9	10	11	12
Per Cent Incidence of Failure	Kraus	76	78	66	69	78	117	91
	Indiana	71.1	53.4	43.4	51.8	56.3	50.2	49.3
Per Cent Test Failures	Kraus	54	55	52	54	58	62	65
Per Cent Flexibility Failures	Indiana	54.1	44.9	36.1	40.8	50.4	44.5	44.9
	Kraus	33	40	40	39	44	40	55
Per Cent Incidence Weakness Failures	Indiana	30.7	32.5	31.1	36.1	45.1	41.4	42.0
	Kraus	39	37	27	30	32	40	36
	Indiana	40.4	20.8	12.3	15.7	11.2	8.8	7.2

*Percentages estimated from graphs in Kraus Report (3).

Apparently an error has been made in graphing the figures for the age 11 group in the Kraus report. The incidence of failure graph shows a tremendous peak at 11 years, estimated at about 117 per cent. Since the sum of the per cent of flexibility failures and the per cent incidence of weakness failures should add up to the total per cent for incidence of failure, it can be seen that an error exists. It is most reasonable to think that the 117 per cent is in error, since the reason for such a large increase in incidence of failure would be rather difficult to explain. However, since it is impossible to determine where the error lies, this age level will be discounted in the comparisons. How the error would affect total figures reported would depend on where the error was made. Other smaller inaccuracies noted can be attributed to errors in plotting, or reading exact figures from the graphs.

Again, it can be noted that the Indiana group is somewhat superior to the Kraus group, in that percentages of failures tend to be lower at all age levels,

except at 6 years. However, in terms of per cent of test failures and per cent of flexibility failures the general trend of failures is very similar for both groups. It is in incidence of Test item failures and incidence of weakness failures that the discrepancies are greatest. The incidence of weakness failures for the Kraus group shows a decided downward trend from age 6 to age 8, but then for some reason takes an upward swing to age 12, while for the Indiana group there is a constant decrease with age, dropping from 40.4 at age 6 to 7.2 at age 12. It is here that it would be important to have a comparison between the two groups in terms of the per cent of weakness failures. Beyond the age of 8, the incidence of test failures is increasing more rapidly than the per cent of test failures, in the Kraus group. This would imply that there is an increase in multiple failures, but not necessarily an increase in per cent of weakness failures.

Kraus and Hirschland state that ". . . children coming into the first grade of the school system are already seriously deficient. Furthermore, it appears that we are unable to alleviate this situation during the time the children are in elementary schools. They leave elementary school in very much the same condition as when they entered it—if anything a little worse." (3, *p. 185*).

It is the intent of this discussion to suggest that if the Kraus figures for per cent of weakness failures were available that a picture more like the one found for the Indiana group might emerge. In that case, it would mean that while there were more children with multiple weakness failures in the Kraus group, there would not be such discrepancies in per cent of weakness failures nor would it be valid to conclude that the children leave the elementary schools in much the same condition as when they entered. They would be increasing in flexibility failures, but decreasing in numbers of children with strength failures.

Owing to the fact that the Indiana group had relatively few multiple failures, their incidence and per cent of weakness failures are very similar.

ANALYSIS BY SEX AND AGE OF THE INDIANA GROUP

The numbers of children, boys and girls, which were measured at each age level are given in Table 4.

TABLE 4

Numbers of Boys and Girls at Each Age Level in the Indiana Group

Sex	Age							Total
	6	7	8	9	10	11	12	
Boys	108	145	124	88	114	121	35	735
Girls	110	138	120	103	110	106	34	721
Total	218	283	244	191	224	227	69	1456

The data in Table 5 reveal the differences in per cents of Test failures that exist between the boys and girls of the Indiana group.

TABLE 5

Number and Per Cent of Test Failures at Different Age Levels for Boys and Girls of the Indiana Group

Sex	Age													
	6		7		8		9		10		11		12	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Boys	64	59.3	75	51.7	58	46.8	49	55.7	60	52.6	60	49.6	17	48.6
Girls	54	49.1	52	37.7	30	25.0	29	28.2	53	48.2	41	38.7	14	41.2

At every age level, the girls show a decided superiority over the boys. Since these tests are designed to measure minimum fitness or to determine the muscular fitness levels of children, we may only conclude that if these tests are valid, our girls are far more muscularly fit than our boys. The natural question is to ask why this should be so. In explaining the differences between the American and European children, Kraus and Hirschland state "The major difference between these two groups is the fact that European children do not have the 'benefit' of a highly mechanized society; they do not use cars, school buses, elevators, or other labor-saving devices. They must walk everywhere—even to school, frequently a long distance. Their recreation is based on the active use of their bodies. In this Country the children are generally conveyed in private cars, or by bus, and they engage in recreation as spectators rather than as participants." (3, *p. 183*) There may be some truth in these statements and it possibly accounts for some part of the differences found to exist between American and European children. But, this reasoning can hardly be applied to the differences observed between boys and girls.

It is felt that the complete breakdown of the figures for each test, by sex and by age, that appears in Table 6 will give at least a partial answer to this question of differences both for the countries and for the sexes.

Reference to Table 6 reveals that at every age level the girls are superior to the boys on the flexibility item, while on all other test items the differences between the two sexes are minor and favor neither sex consistently. Hence, it is largely due to the differences in flexibility of the two sexes that the girls would be termed as more muscularly fit on the basis of the Test results. It is not a new concept that boys are somewhat less flexible than girls, and it does not seem reasonable to characterize the boys as more likely to be health failures or less muscularly fit than the girls, simply because fewer of them can touch their finger tips to the floor with their knees straight.

The facts, that inflexibility increases with age for both sexes and that failures on all of the other items decrease with age, result in a fairly constant per cent of Test failure at all age levels when the sexes and failures on individual items are combined. When these distinct sex and age differences are obliterated by summing them all together, it appears that the children are failing badly at all age levels and that no changes are occurring. For the Indiana group, this is far from the truth. It also might be observed at this point, that in view of these distinct sex and age differences, it hardly seems

reasonable to call any muscular fitness test a minimum standard regardless of age and sex.

TABLE 6

Number and Per Cent of Failures for Each Test Item, by Age and by Sex, for the Indiana Group

Test[1]	Sex	Age 6 n	6 %	7 n	7 %	8 n	8 %	9 n	9 %	10 n	10 %	11 n	11 %	12 n	12 %
F	Boys	44	40.7	61	42.1	52	41.9	45	51.1	58	50.9	57	47.5	16	45.7
	Girls	23	20.9	31	22.5	24	20.0	24	23.3	43	39.1	37	34.9	13	38.2
A+	Boys	11	10.2	5	3.4	1	.8	4	4.5	0	0	3	2.5	0	0
	Girls	14	12.7	10	7.2	1	.8	4	3.8	0	0	1	.9	1	2.9
A−	Boys	38	35.2	28	19.3	10	8.1	11	12.5	6	5.3	9	7.5	1	2.9
	Girls	36	32.7	25	18.1	11	9.2	14	13.6	13	11.8	5	4.7	2	5.9
P	Boys	3	2.8	4	2.8	3	2.4	2	2.3	1	.9	3	2.5	1	2.9
	Girls	4	3.6	2	1.4	3	2.5	2	1.9	5	4.5	3	2.8	1	2.9
UB	Boys	2	1.9	0	0	0	0	0	0	0	0	0	0	0	0
	Girls	2	1.8	0	0	1	.8	0	0	0	0	0	0	0	0
LB	Boys	0	0	0	0	1	.8	0	0	0	0	0	0	0	0
	Girls	3	2.7	0	0	1	.8	1	1.0	0	0	0	0	0	0

[1] F — flexibility; A+ — abdominals plus psoas; A− — abdominals minus psoas; P — psoas; UB — upper back; LB — lower back.

The greatest per cent of failure on the strength items is on the abdominals minus psoas and it appears to be very valuable in detecting weaknesses in the muscle groups concerned. It is heartening to note, however, the very rapid decreases occurring in both sexes as the age increases—the boys decreasing from 35.2 per cent failure at age 6 to 2.9 per cent at age 12, while the girls' failure rate progresses from 32.7 per cent failure at age 6 to 5.9 per cent failure at age 12.

The psoas test results show only small percentages of failure for both sexes at all ages. The average failure for the boys is 2.3 per cent and for the girls 2.8 per cent.

In the entire population of 1,456 children, there were only 5 failures on the upper back test, 6 failures on the lower back test, and none after nine years of age on either test. It would appear that, for the Indiana group at least, these items contribute little to the Test and could be eliminated without noticeably affecting the results. If back strength is to be measured, as part of a fitness test, other more discriminating items need to be substituted.

The figures of Table 6 indicate that after the Indiana boys and girls reach 11 years of age there is no strength item on which they show as much as 8 per cent failure. At this age, these boys and girls are coming very close to the European children in the strength items. The major discrepancy in Test failure must be attributed to differences in the flexibility item. The cause of

this difference in flexibility, as measured by ability to touch the finger tips to the floor with knees straight, and its significance to muscular fitness are problems needing further exploration.

Summary and Conclusions

The Kraus-Weber Test of muscular fitness was administered to a population of 1,456 elementary school children from an Indiana city. Repeats of the test were given to 215 children, and the grip strength of the preferred hand was measured for 126 of the children.

The Test and the individual items of the Test were all found to be highly reliable as evidenced by coefficients exceeding .950 in all cases.

No relationship appeared to exist between the Kraus-Weber Test and the Grip Strength Test. The differences between the mean grip strength of children passing the Kraus-Weber Test and those failing this test lacked statistical significance. This was true for both boys and girls, although a significant difference between the sexes was found, in favor of the boys.

The Indiana group was found to be somewhat superior to the Kraus group in all failure comparisons. However, the per cent of test failures and per cent of flexibility failures at the different age levels showed very similar trends for the two groups.

The major discrepancies between the two groups were in incidence of Test failure and incidence of weakness failure.

The analysis of the Indiana results by age and sex revealed the following:

The girls are substantially superior to the boys at all age levels on the Kraus-Weber Test. This superiority is due to their greater success on the flexibility item at all age levels.

For both sexes, inflexibility increases with age, with more pronounced increases for the girls at the higher age levels. As the children become older, the flexibility measures of the sexes are coming closer together.

For both sexes, there is a decided decrease in strength item failure as age increases. By the time the children are 11 years of age, the per cent of failure on any strength item is less than 8 per cent. Their strength measures approach very closely those of the European children at this age.

When flexibility and strength weakness figures are added together, without regard for sex, differences which exist are obliterated, and the results imply that no changes are occurring with age, in the children. Hence, these figures should be separated in order to reveal the true failure picture.

Since not more than six children in the entire population of 1,456 failed on each of the back strength items, these items are not discriminatory and should either be eliminated or some other back strength tests substituted.

The discrepancy in the flexibility of the Indiana children and the European children, as well as the discrepancy between the boys and girls in flexibility are not readily explained. There is need for further exploration of the reason for these differences and the significance of flexibility failure to muscular fitness.

REFERENCES

1. GUILFORD, J. P. *Fundamental Statistics in Psychology and Education.* New York: McGraw Hill Book Co., Inc., 1942.
2. KRAUS, HANS, M.D., and HIRSCHLAND, RUTH P. Muscular fitness and health. *Journal of the American Association for Health, Physical Education, and Receation,* **24**: 17, Dec. 1953.
3. ———. Minimum muscular fitness tests in school children. *Research Quarterly,* **25**: 178-188, May 1954.
4. LARSON, L. A. A factor and validity analysis of strength variables and tests with a test combination of chinning, dipping, and vertical jump. *Research Quarterly,* **11**: 82-96, Dec. 1940.
5. ———. A factor analysis of motor ability variables and tests, with tests for college men. *Research Quarterly,* **12**: 499-517, Oct. 1941.
6. MCCORMICK, T. C. *Elementary Social Statistics.* New York: McGraw Hill Book Co., Inc., 1941.
7. PHILLIPS, MARJORIE. Study of a series of physical education tests by factor analysis. *Research Quarterly,* **20**: 60-71, March 1949.
8. STANSBURY, EDGAR. A simplified method of classifying junior and senior high school boys into homogeneous groups for physical education activities. *Research Quarterly,* **12**: 765-776, Dec. 1941.

(Submitted 11/1/54)

Measurement of Kinesthesis[1]

M. GLADYS SCOTT
University of Iowa
Iowa City, Iowa

Abstract

This is an attempt to establish tests for measurement of kinesthesis. For the first sample of 100 college women, 28 measures of kinesthesis and 2 of motor ability were administered. For the second sample of 70 college women, 15 of the previous tests were given and one new one was added. Analysis was made of the quality of the test items and of the inter-relationship of tests. Most of the tests were adequate in reliability. No single test was sufficiently highly related to the criteria employed to be used alone as a measure of kinesthesis, though several combinations gave fair validity. Because of this fact and because of the very low number of significant intercorrelations, it was concluded that kinesthesis is composed of a series of specific functions.

IN SPITE OF THE FACT that physical educators have recognized for years that the sense of kinesthesis must have relationship to motor performance, we really know very little about how to identify the varying degrees of this sensory acuity. Efforts at measurement have been sporadic and not very fruitful. An attempt at definition and identity, which is essential as a starting point in measurement, is impossible because the facts are lacking. It represents a challenging type of investigation, but in analogy is a situation of trying to lift one's self by the boots without even the boot straps to grasp.

This investigator is not the only one seeking an answer to these issues. For the period preceding 1950, there are studies such as Bass (1), Young (25), Fisher (4), Phillips (14), Seashore (20), Honzik (6). Since 1950, they have become more numerous; there have been studies such as Stevens (21), Wiebe (23), Witte (24), Roloff (18), Henry (7), Espenschade (3), Riley (17), Garrison (5), Worchel (22), Mumby (12), Norrie (13), Lindsay (9), Meday (10), Mullin (11), Rehling (16). Studies are now coming not from two or three centers but from many different ones scattered from coast to coast.

The earlier studies aimed at validating empirically selected test items, and they have opened the way for the later studies of factor analysis, correlation with learning rate, effects of practice, relationship to specific skills, and measurement of growth or maturation through childhood and youth. Why do more and more people devote time and effort to work on what seems to be an intangible? Because it is believed to hold at least a partial answer to the secret of individual differences in motor accomplishment and in learning rate. It can give the physical educator a tool with which to instruct and serve as an aid to understanding, guiding and motivating individuals. Method might be

[1]Based on papers presented at the International Congress on Essentials of Physical Education for Youth, Connecticut Valley Colleges, April 13, 1954; and at the Research Section of the Central Association for Health, Physical Education, and Recreation Convention, Lincoln, April 9, 1954.

so improved as to lead to considerably greater accomplishments, and development of an ability in the learner to be more independent and more capable of self-directing his learning in the future. This at least is the faith on which this investigator continues, and is probably similar to that of others.

Probably the most forceful statement of the significance of kinesthesis in motor performance comes from Ragsdale (15) in a discussion of motor learning. He makes statements such as "kinesthetic perception is fundamental but must be interpreted and acted upon," "blindfold practice makes for greater dependence on kinesthetic cues," "manual guidance may help to develop kinesthetic perception of the activity." He indicates many practical ways in which kinesthesis may be utilized to facilitate learning. It is faith in such a use that is bringing more and more efforts at relating kinesthesis to learning, measurement of the trait, and increased understanding of the ways in which it functions in motor behavior.

Background of the Study

The aim of these investigations is to establish a practical and valid method of determining kinesthetic function with a view to understanding better the learning process. The successive stages of work have required the following sub-problems.

A. Identification of ways in which kinesthesis is demonstrated in motor performances
B. Selection of tasks, based on these expressions of kinesthesis, suited to objective measurement
C. Determination of the inter-relationship of measures and the contributions of each measure to the concept of total kinesthesis
D. Ascertaining the relationship of levels of kinesthesic acuity to learning facility and motor accomplishment

Because of the nature of the problem and the lack of clear-cut evidence, the work has been characterized by replication, new hypotheses, gradual evolvement of new and revised forms of measurement. The data of this report are limited mainly to one phase of the work, with citation of a few points from earlier, unpublished studies.

Kinesthesis has been recognized by both psychologists and physiologists as a sensory property of the human being with implications for motor accuracy and precision, individual differences, empathic appreciation, and even determination of neural normalcy. Kinesthesis is referred to frequently as the "motor sense" and for this investigator is defined as the sense which enables us to determine the position of segments of the body, their rate, extent, and direction of movement, and the position of the entire body, and the characteristics of total body motion. Stimuli pertinent to this sensation are carried by special receptors and nerves just as the stimuli of vision are carried by neural pathways peculiar to the eye, or the stimuli for smelling by neurons specific to the sensation of smell.

The receptors of kinesthesis are identified by the physiologist (2) as a) the muscle spindles, around the muscle fibers; b) the Golgi corpuscles in the tendons; c) the Pacinian corpuscles in tendons and articular cartilage of joints; and d) free nerve endings in the muscles, tendons and joints. They are stimulated by tension or pressure. The labyrinthine source of stimulation is activated by movements of the head when the head moves independently or in unison with the total. The vertical canals apparently function in antero-posterior or lateral movements while the horizontals are most helpful in the twisting, spinning movements.

The eyes aid indirectly in contributing what might be called "orienting information" such as location of the horizontal, distances, changes in relative position of the organism, and movement of persons and objects about the observer.

On this background of psychological-physiological perspective the various proposed tests have been developed with the usual assumption of face validity. The number of tasks that could be selected are almost limitless. Some do not lend themselves to objective measurement, some are not very interesting to subjects, some require elaborate equipment. If this is to be practical for the teacher, possible to use in the gymnasium as well as in the research laboratory, then special apparatus must be eliminated as completely as possible.

Procedure

Part A. Twenty-eight tasks were selected as meeting the standards of testing and as having face validity. Face validity, as used here, means that by hypothetical concepts based on present information concerning kinesthesis the tasks would appear to require some kinesthetic functioning. The tests were selected with the idea of sampling different phases of kinesthetic perception. This was based on a general study of the literature, previous experience with testing of this type, and results of a factor analysis study recently completed (12). These component phases are ability to repeat muscle contractions with a force identical to that which one has just exerted; ability to put the arms, legs and trunk in positions prescribed by visual or oral cues; balance and weight control; manipulative precision with the hand; orientation in space; and the ability to imitate promptly a simple co-ordination which has been demonstrated, i.e., to translate empathic understanding into concept and thence to movement. To this series, broad jump (maximum) and motor ability were added as measures of accomplishment.

TEST DESCRIPTIONS[2]

1. *Leg force*—Required to extend leg with exactly 20 lb. of force. Score = deviation from 20 lb.
2. *Arm force*—Required to extend arm with exactly 10 lb. of force. Score = deviation from 10 lb.

[2]Detailed test instructions may be obtained from the author. The abbreviated statement is given here to give an idea of the task involved in the test item.

3. *Pull*—Required to pull with both arms exactly 10 lb. Score = deviation from 10 lb.
4. *Push*—Required to push with both arms exactly 10 lb. Score = deviation from 10 lb.
5. *Balance leap*—Sideways leap, forward bend, moving object on floor and subsequently holding balance at least 5 sec. Score = time balance is held.
6. *Balance stick*—Balance on one foot placed lengthwise on 1 in. square block, eyes closed. Score = total time.
7. *Weight shifting*—Weight on one foot on solid block, shift weight toward other foot on scale until ½ weight is on each foot. Score = deviation from prescribed amount.
8. *Body sway*—Center of gravity apparatus, standing on one foot sideways to scale, eyes closed, for 10 sec. Score = range registered on scale.
9. *Leg raising*—Side lying abduct leg according to model shown (20°). Score = deviation from 20°.
10. *Thigh raising*—Standing, raise leg, knee bent, till thigh is horizontal. Score = deviation from horizontal.
11. *Arm raising*—Standing, abduct arms to horizontal. Score = deviation from 90°.
12. *Wrist extension*—Forearm supported on table, hand relaxed over edge. Extend wrist according to model shown (30°). Score = deviation from 30°.
13. *Finger spread*—Forearm supported on table, fist closed, raise hand but not elbow and touch thumb and little finger on targets. Score = deviation from target center.
14. *Lower leg flexion*—Standing on one foot, knee pointing downward, flex lower leg of free foot until lower leg is horizontal (90°). Score = deviation from 90°.
15. *Wrist flexion*—Forearm supported on table, hand relaxed over edge. Extend wrist according to model (20°) (flexion). Score = deviation from 20°.
16. *Ball balance*—Golf ball in small trough. Keep rolling constantly, crossing finish line but not contacting end of trough. Score = errors made in 30 sec.
17. *Target pointing*—Standing sideways to target, pointer in hand, eyes closed. Raise pointer to target. Score = deviation from center of target.
18. *Horizontal lines*—With eyes closed, draw 4 horizontal lines (parallel to one observed before closing eyes): (1) left to right with hand, (2) right to left with hand, (3) left to right with foot, (4) right to left with foot. Score = deviation from horizontal.
19. *Walk and turn*—Facing down a straight line. With eyes closed, step right, and turn right, step left and turn left, repeat. Score = angle of deviation from the straight line.
20. *Broad jump specified*—Stand toeing a line. With eyes closed jump so feet are just across a line 24 in. ahead. Score = deviation of heels from a 24-in. line.
21. *Floor target*—Look at targets on floor and with eyes closed touch index finger to target point. Score = deviation of heels from a 24 in. line.
22. *Broad jump maximum*—From take-off board, jump forward as far as possible. Score = distance from take-off to heel.
23. *Motor ability*—Weighted sum of scores on broad jump (maximum), obstacle race and basketball throw for distance (19).
24. *Arm swinging*—Four-count arm exercise, swinging forward and sideways, with alternate arm holding first count. Score = rating on precision and number of trials required.
25. *Arm circling*—Arm circling in opposite directions. Score = rating on precision and number of trials required.

The subjects were 100 college women selected at random. Their ages varied from 17 to 35 years, but were predominately in the 18 to 20 year age range.

Reliability of tests with several trials was computed on the alternate halves and stepped-up with the Spearman-Brown formula. The tests with only one

trial were administered to part or all the subjects later in the testing hour; on other tests the subjects were asked to come back later for repetition of the tests. Thus, the number on which reliability was computed varied from 45 to 100. The reliability was low on several tests. The five lowest were dropped from further consideration in this analysis. Intercorrelation for all the remaining items were determined. A 5-per-cent level of confidence was adopted, for the 100 cases requiring an r of .195.

The validity was computed by correlating with a battery recommended by Young (25) and with another recommended by Roloff (18). Formulas for these two batteries are:

Sum of T-scores on Balance Stick + Leg Raising + Arm Raising (Young)
.75 Balance Stick + 4.7 Arm Circling — Arm Raising — Weight Shifting + 50 (Roloff)

Finally, multiple correlations were computed to estimate the validity of different combinations of tests.

Part B. Subsequent to the analyzing of data obtained on Group A as outlined above, data were collected on another comparable sample of 70 college women. Some of the tests were retained in the same form, trials were increased on some in an attempt to improve reliability; and one new test was added—triangle walk. This was a test recommended by Worchel (22). He says that for movement through short distances, kinesthetic cues are responsible for keeping the subjects oriented and that those with defective semicircular canal function did more poorly on this type of test than those with a normal labyrinth. The total number of tests given was 18. As in the previous studies, the subjects were tested individually, and were either blindfolded or had eyes closed.

Computation was made of reliability of each test. Two were not considered further because of low reliability coefficients. The remaining 16 were used for intercorrelation and validity computations. As in the previous study, the Young (25) and Roloff (18) criteria were used. A third criterion was also used, one obtained from data on Group A of this report.

Analysis of Data

Part A. Table 1 contains the reliability of the tests for both groups. These compare favorably in those items which have been used previously. For the majority of tests used with both A and B groups, the coefficients obtained on the two samples are similar.

Intercorrelations of all items showed very few significantly greater than zero, that is, unaccounted for by chance. These coefficients are more meaningful in smaller groups of items which had been put into the experimental battery on the hypothesis of similarity of kinesthetic function. Therefore, the intercorrelations are presented in relation to the hypotheses on which the trial battery was built. Analysis of these hyoptheses follows on data from group A.

TABLE 1

Reliability of Kinesthesis Measures

Test Item	Group A r	Group A N	Group B 70(a) r	Previous Studies Roloff 200(a) r	Previous Studies Scott(b) 104(a) r
Leg force	.64[c]	100			
Arm force	.79[c]	100	.50[cd]		
Pull	.63[c]	45			
Push	.83[c]	100	.54[cd]		
Balance leap	.68[c]	100	.71[c]	.65	.82
Balance stick	.86[c]	100	.83[c]	.86	.72
Weight shifting	.37	50	.35		.68
Body sway (side)	.63	100	.72		
Leg raising	.93[c]	100	.76[cd]	.94	.98
Thigh raising	.64	49			
Arm raising	.91[c]	100	.81[cd]	.80	.80
Wrist extension	.68[c]	94			
Finger spread	.85[c]	100	.78[c]		
Lower leg flexion	.82	49	.85		
Wrist flexion	.51[c]	94			
Ball balance	.81[c]	94	.88[c]		
Target pointing (pointer)	.72[c]	69	.77[c]		
Horizontal lines	.52	50	.51[c]		
Walk and turn (angle)	.43	71	.65		
Broad jump (specified)	.27	76	.36		
Floor target	.57[c]	100			.73
Arm swinging					.78
Arm circling					.64
Vertical lines	.00	50			
Body sway (forward)	.39	100			
Target pointing (finger)	.15	69			
Trunk horizontal	.11	49			
Walk and turn (distance)	.21	31			
Triangle walk			.52		

a Number of cases on which coefficients were computed.
b Unpublished study.
c Computed on alternate halves and stepped up with Spearman-Brown formula.
d Difference between r of A and B group is significant at the 5% level.

A. *Muscular contraction of a known amount is a function of kinesthesis and each subject should be equally accurate with various muscle groups.*

The items which would be expected to be related on this basis are listed in the first part of Table 2. Other tests were related significantly to one or more of those listed in Table 2 and these coefficients are also shown.

TABLE 2

Correlations Between Measures of Specified Amounts of Contraction and Others Which Show Significant Correlation with Them

Measure	1	2	3	4	5
1. Leg force					
2. Arm force	.02				
3. Pull	.05	.30			
4. Push	−.12	.19	.22		
5. Broad jump (spec)	−.25	.06	.01	.02	
Arm swinging	.24				
Ball balance		.22			
Balance leap		.20			
Balance stick				−.21	
Horizontal lines				.24	
Floor target				−.26	
Finger spread					.22
Weight shifting					.20
Broad jump (max)	.18				−.09

Interpretation:

1. The use of the arms in adapting effort appears to be inter-related in these tasks.

2. The use of the arms in adapting effort is apparently unrelated to similar function in the legs.

3. Facility in adapting effort is general enough to affect performance in a variety of tasks, apparently facilitating control of the hands and arms.

4. The inverse relationship of balance stick and floor target with the push is not readily apparent.

B. *Balance ability is a specific function of kinesthesis.*

Four tests were judged to measure some aspect of balance. Correlations of other tests with the balance items are so uniformly significant that they are presented also in Table 3. (See facing page.)

Interpretation:

1. Three of the four tasks selected have some inter-relationship, though less than might be anticipated. Balance stick and body sway are in a stationary position while balance leap involves movement. This may partially explain the rather low r's.

2. The other items with which they are related represent in general complex skills, with relatively little balance, apparently seeming to be general co-ordination instead.

3. The relationship with the arm movement is not very meaningful. It may be that the arm action in balance is a very significant one, not just a mechanical asset as sometimes is assumed in analysis of skills.

TABLE 3

Correlations Between Measures of Balance and Others Which Show Significant Correlations with Them

Measure	1	2	3
1. Balance leap			
2. Balance stick	.40		
3. Weight shifting	−.04	.14	
4. Body sway	.12	.23	−.01
Broad jump (spec)			.20
Arm force	.20		
Broad jump (max)	.38	.50	
Motor ability	.43	.50	
Arm swing	.20	.29	.21
Arm circling	.29	.23	.19
Arm raising			.20

C. *Ability to assume and identify body position is a general capacity operating throughout the body.*

TABLE 4

Correlations Between Measures of Ability To Assume and Identify Body Position

Measure	1	2	3	4	5	6
1. Leg raising						
2. Thigh raising	−.08					
3. Arm raising	−.27	−.11				
4. Wrist extension	−.09	−.09	.04			
5. Wrist flexion	−.05	.02	−.06	.11		
6. Lower leg flexion	.12	−.10	−.06	.09	−.02	
7. Finger spread	−.06	.16	.01	.09	−.16	−.02

The only other items related to this group are arm circling and arm swinging, related to leg raising (.32) and arm raising (.19), respectively.

Interpretation:

1. The hypothesis is completely untenable. Either the ability to identify position of the various segments is specific for each segment or some factor, such as variations in experience, leaves these subjects using their potentialities in very different ways in the different segments.

2. While all the coefficients except one are insignificant, the majority are negative. This is in disagreement with previous studies such as Young (25) and Roloff (18). It is probably a chance occurrence.

D. *Precision in use of the hands is a basic function in manipulative skills and is made possible by perception of hand and arm position and movement.*

The four items selected on this hypothesis failed to show any relationship. However, there were a number of other tests which were related to them. (See Table 5.)

TABLE 5

Correlations Between Measures of Manipulative Hand Function and Others Which Show Significant Correlation with Them

Measure	1	2	3	4
1. Ball bounce				
2. Target pointing (pointer)	.08			
3. Horizontal lines	.17	.00		
4. Finger spread	—.10	.08	.11	
Arm force	.22			
Arm swinging	.19			
Arm circling	.29		.23	
Push			.24	
Broad jump (maximum)				—.19
Broad jump (specified)				.22

Interpretation:

1. These particular tasks are unrelated and the hypothesis of common function in these tasks is untenable.

2. Ball balance and horizontal lines are related to arm functioning of various types and are understandable as such. Ball balance apparently is more an arm function than one of the hands.

3. The relationship between finger spread and broad jump events is unaccountable.

4. The lack of relationship of these items and wrist flexion and extension lends further support to the probability of specificity of functions.

E. *Orientation of the body in space facilitates precise movement and ability in various types of tasks involving space orientation should be interrelated.*

Table 6 shows the four tests to be unrelated among themselves, though four other items are related to them. (See facing page.)

Interpretation:

1. The hypothesis is untenable; these tasks are apparently specific.

2. The weight shifting and finger spread might be considered a form of space orientation; similarity between specified broad jump and weight shifting is logical, that with finger spread is not and seems to contradict the concept of specificity of functions. Worchel (22) tested ability of subjects both with normal and insensitive labyrinths. He says that for long distances

TABLE 6

Correlations Between Measures of Orientation in Space and Others Which Show Significant Correlations with Them

Measure	1	2	3
1. Walk and turn			
2. Broad jump (specified)	—.00		
3. Floor target	—.08	.05	
4. Target pointing	.04	—.13	—.08
Leg force		—.25	
Weight shifting		.20	
Finger spread		.22	
Push			—.26

spatial imagery enters into orientation of the subject in moving in a specified angular path, but that on short distances, kinesthetic cues are mainly responsible. He further stated that those with functional semi-circular canals did better on the tests than those without this function. All the tasks in this study would appear to be concerned with short distances. They do involve very different types of movement, however.

F. *Learning of a new skill is facilitated by kinesthetic cues which make for similarity of achievement in tasks to be learned.*

The total group of items shown in Table 7 represents what would appear to be very diversified capacities.

TABLE 7

Correlations Between Measures of Learning of New Skills and Others Which Show Significant Correlations with Them

Measure	1	2	3	4
1. Arm swinging				
2. Arm circling	.42			
3. Broad jump (maximum)	.26	.24		
4. Motor ability	.30	.31	.84	
Leg force	.24			
Balance leap	.20	.29	.38	.43
Balance stick	.29	.23	.50	.50
Weight shifting	.21	.19		
Arm raising	.19			
Ball balance	.19	.29		
Leg raising		.32		
Horizontal lines		.23		
Finger spread			—.19	

Interpretation:

1. The hypothesis seems extremely plausible. It is the only assortment of hypothetically related items in which all are significantly related.

2. Learning rate is related to accomplishment level.

3. Whatever the source of varitation in learning rate is, it operates in many other tasks established in this study. The spread of relationship to other tasks, and the uniformity of positive relationship (only one negative) lend further support to the belief that previous experience is a contributing factor in learning and adapting to new tasks. Since these items were all selected on face validity, i.e., involving kinesthetic perception, one might conclude that kinesthetic perception is the connecting link between previous experience and learning of a new motor task.

In considering intercorrelations of all items (see Table 10, p. 338), it is apparent that the relationships are in general low. They are lower than obtained in previous studies, and the question is raised as to the performance of the group and the test reliability. Table 1 showed the reliability to be not markedly different where comparison was possible on identical tests. The range and variance were found to be similar and in most cases large. Subjectively, it is believed that the subjects were not as uniformly interested and motivated as they might have been and as previous groups may have been.

The face validity of the individual items was considered reasonable. No exact and completely verified criterion exists for statistical computation of validity. However, two of those which have been used previously were used in this study. The one was established by Young (25) and the second by Roloff (18). These results appear in Table 8.

It is apparent first of all that very few of the tests are significantly greater than zero (.19 at the 5% level). For instance, on the Young battery only balance leap, arm swinging, and arm circling are significant, with horizontal lines, body sway, and leg force approaching significance. For the Roloff battery the only significant r's are on balance leap, arm swinging, horizontal lines, and walk and turn, with ball balance approaching significance.

The multiple R's and best eight combinations are presented in Table 9. The size of the coefficients for the first six are practically identical. There is little basis in the size of the R's for choice between them. From the standpoint of administrative speed, the arm swinging and balance leap would appear optimum. The third, fourth, or fifth test does not add enough to pay for time consumed according to present evidence.

Part B. Reliability of each item is given in Table 1. In comparing these coefficients with the ones on the earlier study, it will be noted that four of the coefficients dropped to the extent of a statistically significant difference. These four are arm force, push, leg raising, and arm raising. In two of these tests there have been several comparable samples on which reliability has been computed. In one case, this is lower than any of the others, viz., leg raising

TABLE 8

Validity of Tests on Group A and B as Compared with the Young, Roloff and Scott Criteria

Test Item	Criteria				
	Young		Roloff		Scott[a]
	A	B	A	B	B
Leg force	.15		.11		
Arm force	−.01	.07	.13	.00	.05
Pull	−.05		.00		
Push	−.03	−.07	−.01	−.08	.25
Balance leap	.46	.13	.31	.17	
Balance stick					.24
Weight shifting	.11				
Body sway	.16	.00	.09	.33	.43
Leg raising		.36	−.02	−.04	.12
Arm raising					.14
Thigh raising	−.07		.00		
Wrist extension	.02		.04		
Finger spread	−.02	.06	−.07	.13	−.08
Lower leg flexion	.13	.11	.14	.05	.03
Wrist flexion	−.08		−.11		
Ball balance	.02	.11	.18	.07	.25
Target pointing (pointer)	.08	−.05	.14	.09	−.02
Horizontal lines	.19	.05	.23	.02	
Walk and turn	.08	.10	.20	.10	−.08
Triangle walk		.09		.15	.21
Floor target	.04		.00		
Broad jump (specified)	−.04		.06		
Broad jump (maximum)	.38		.40		
Motor ability	.41		.36		
Arm swinging	.44	.14	.47	.49	
Arm circling	.40	.08			.44

[a] Scott criterion = 19 arm swinging + 8 balance leap − horizontal lines.

.76 as compared with .93, .94, and .98. For arm raising the drop is probably to more nearly the true reliability to be expected—.81 as compared with .91, .80 and .80 in other studies. In the other two tests there are only the two samples for comparison.

The validity values on the three criteria are in Table 8. The coefficient was not computed if the test was a part of a battery. A few points seem to be apparent from study of data from various samples.

1. Many of the tests agree fairly well from sample to sample and from one criterion to another in terms of level of relationship.
2. Negative coefficients are without exception insignificant.
3. There is more agreement from one criterion to another within samples than between samples for each criterion.

4. The tests which correlate significantly with the criterion taken from Part A are push, balance stick, body sway, ball balance, triangle walk, and arm circling. This is a larger proportion of valid items than with either the Young or Roloff criterion for either this sample or the previous one. This may mean simply that the test array is more homogeneous and it may mean that the criterion is a more adequate one.

TABLE 9

Validity of Best Test Batteries

Battery	R
COMPUTED ON DATA FROM GROUP A	
Young criterion:	
1) 7.0 arm swinging + 3.0 balance leap — .4 horizontal lines — 1.0 body sway	.60
2) 19.0 arm swinging + 8.0 balance leap — 1.0 horizontal lines	.59
3) 7.0 arm swinging + 2.5 balance leap — 1.0 body sway	.59
4) 5.0 arm swinging + 2.0 balance leap	.58
Roloff criterion:	
1) 4.8 arm swinging + 1.0 balance leap — .4 walk and turn — .3 horizontal lines — .5 ball balance	.58
2) 16.0 arm swinging + 3.0 balance leap — 1.0 walk and turn — 1.0 horizontal lines	.58
3) 16.0 arm swinging + 3.0 balance leap — 1.0 walk and turn	.55
4) 17.0 arm swinging — 1.0 walk and turn	.51
COMPUTED ON DATA FROM GROUP B	
Scott criterion:	
1) 4.0 arm circling — .9 push — 1.0 ball balance — .7 body sway	.66
2) 6.0 arm circling + .9 balance stick — .6 ball balance — 1.0 body sway	.63
3) 100 — .8 triangle walk — 1.3 ball balance — 1.0 body sway	.55
4) 6.0 arm circling — 1.0 body sway	.52
5) 5.0 arm circling — 1.0 push	.51

Table 11 (p. 340) gives intercorrelations of tests for Group B. Of the 120 correlations only 21 are significant. This may lend support to the idea of specificity of functions involved. Very few of the significant r's are negative although there are many negative r's in the table. The proportion of negatives is smaller than for the correlations obtained in Part A. It seems probable that many of the negative ones are the result of sampling variation.

Multiple correlations were determined with various combinations of items found valid relative to the criterion from Part A. These appear in Table 9. The R's on the shorter batteries are comparable to those correlated with the other criteria, but the longer batteries are slightly higher.

Discussion and Observations

Of practical significance, perhaps, is the fact that the tests involved in these new batteries are all more interesting to the subjects than were the Young items. This is a great help in getting full co-operation from the subjects.

While these results are not as conclusive as one might hope for, there are certain things which seem to stand out from these data.

1. A balance test shows up favorably in every study that has been made with a balance test included. Whether this means that the neural mechanism of balance is similar to that used in other test items or simply that acuity of labyrinthine function parallels proprioceptor function cannot be answered at this time. However, it seems advisable to put a balance test into any kinesthesis battery. The two used here have been consistently reliable and fairly valid.

2. Arm swinging checks with both of these batteries as it did previously with a longer criterion of many test items. Its consistency is apparent. Analysis of the test seems to indicate two capacities involved, i.e., ability to perceive kinesthetic cues and to execute precise movement in plane, distance and timing. It might be considered a learning capacity, but on the other hand, it is more than perception of the problem. Subjects discussing the test afterward sometimes said they knew what they were trying to do, but they had not been able to execute the movements, becoming lost in the sequence, or were inaccurate in movement. Arm circling has similar characteristics, correlating .42 with arm swinging and having similar validity coefficients whenever we have used them.

3. The ball balance and horizontal lines are arm manipulations and though they intercorrelate only .17, they might show up better comparatively if the trials were increased and reliabilities raised, especially the horizontal lines.

4. The walk and turn, leg force, and body sway also need work toward improving the reliability.

The correlations of motor ability and broad jump (which is a part of this MA battery) are reasonably high for each of the two criteria of kinesthesis and also with the balance tests. They approach significance on leg force, target pointing (with pointer), floor target, but appear to be inversely related to finger spread. Correlations do not seem to be high enough to indicate that these are learned tasks. It can be assumed that it is some capacity related both to learned skills and adjustments to new motor acts. At the present time it is hypothesized on the basis of experience in this area of testing that kinesthetic perception could be the common element.

Actually, this question of relationship between kinesthetic functioning and motor achievement is the reason for all of our efforts pertaining to kinesthesis. Practical application is the ultimate objective in obtaining information of this type. It is hoped it will lead eventually to better understanding of the process of motor learning, of individual differences in motor learning, of

TABLE 10

Intercorrelation of Kinesthesis Tests on Group A

Test Item	1	2	3	4	5	6	7	8	9	10
1. Leg force										
2. Arm force	.02									
3. Pull	.05	.30								
4. Push	−.12	.19	.22							
5. Balance leap	.03	.20	−.03	−.04						
6. Balance stick	.07	.05	.02	−.21	.40					
7. Weight shifting	.06	.17	−.02	−.01	−.04	.14				
8. Body sway (side)	−.13	.01	−.03	−.04	.12	.23	−.01			
9. Leg raising	−.02	−.00	−.08	−.05	.18	−.01	−.06	.14		
10. Thigh raising	.06	.07	−.01	−.08	−.08	.07	.14	.14	−.08	
11. Arm raising	.15	−.02	−.03	.18	.00	−.00	.20	−.15	−.27	−.11
12. Wrist extension	.03	−.09	−.17	−.10	.12	.10	.04	−.12	−.09	−.09
13. Finger spread	−.15	−.14	−.13	−.07	.06	−.02	.03	−.10	−.06	.16
14. Lower Leg flexion	−.12	−.07	−.05	−.09	.03	.11	.13	.03	.12	−.10
15. Wrist flexion	.08	−.00	−.00	.10	−.01	−.13	−.12	.02	−.05	.02
16. Ball balance	.15	.22	−.02	.17	.06	.05	.14	.04	.03	.07
17. Target pointing	.17	−.04	−.06	.15	.16	.03	.09	−.12	−.02	.00
18. Horizontal lines	−.05	.14	.07	.24	.01	.18	.03	.06	.05	.13
19. Walk and turn	−.10	.04	−.02	.09	.03	.12	.08	−.04	.03	−.02
20. Broad jump (spec)	−.25	.06	.01	.02	−.01	−.10	.20	−.04	−.01	.13
21. Flood target	−.11	−.09	−.02	−.26	.03	.12	−.04	.05	−.01	.18
22. Broad jump (max.)	.18	−.01	−.09	−.10	.38	.50	.09	.03	.01	−.08
23. Motor ability	.08	.04	−.01	−.14	.43	.50	.05	.04	.14	−.01
24. Arm swinging	.24	.11	−.08	−.06	.20	.29	.21	.05	.13	.03
25. Arm circling	−.02	.10	.03	.08	.29	.23	.19	.02	.32	−.05

new methods to employ this motor sense in the learning process, and the maturation of the capacity for meaningful perception.

There are implications in these data which would appear to give direction to further studies. With the mounting evidence toward the indication that the kinesthetic sensation is a complex of many independent and specific capacities or component elements, certain steps are indicated.

1. Analysis must be made to determine those components.
2. Test items will need to be created for many or all of the components.
3. Test batteries will doubtless be fairly large when developed ultimately in order to adequately understand the individual variations.
4. Determination of relationship between kinesthetic acuity and motor learning will be more complex. Achievement in one activity may be related to certain elements and success in another activity more related to other aspects of kinesthesis.

Conclusions

In light of data obtained from these two groups of 100 and 70 cases respectively the following conclusions are drawn.

TABLE 10 (*Continued*)

11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
.04														
.01	.09													
−.06	.09	−.02												
−.06	.11	−.16	−.02											
−.04	.04	−.10	−.03	.03										
.13	.05	.08	−.04	.09	.08									
.07	−.01	.11	−.03	.01	.17	.00								
.00	.03	.06	.12	.01	−.13	.04	.00							
.01	−.02	.22	.07	.08	.04	−.13	−.04	−.00						
−.04	.08	.03	.12	−.06	−.05	−.08	.01	−.08	.05					
.03	.16	−.19	−.01	.11	.09	.18	.06	−.01	−.09	.15				
−.03	.09	−.17	.05	.07	.07	.14	.07	.02	−.08	.17	.84			
.19	.14	−.07	−.16	−.04	.19	−.05	.13	.00	.06	.02	.26	.30		
.11	−.14	−.13	.11	.04	.29	.12	.23	.10	.12	−.07	.24	.31	.42	

1. Several tests constructed on the basis of face validity for kinesthesis yielded reasonably reliable results which are consistent from one sample to another.

2. The tests in general show little inter-relationship. This would lead one at present to assume considerable specificity of function.

3. There is no single test which has adequate validity compared with the criteria available to date.

4. The level of validity of a test item depends upon the criterion being used.

5. Several combinations of tests appear to have a validity value satisfactory for further use.

6. The sensation of kinesthesis is made up of many elements or forms of response. There is little or no evidence that it might be a general capacity.

7. Much more work needs to be done toward refinement of the measures.

TABLE 11

Intercorrelation of Kinesthesis Tests on Group B

Test Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Arm force															
2. Push	.26														
3. Target pointing	−.11	−.02													
4. Walk and turn	−.11	−.20	−.02												
5. Balance leap	.05	.02	−.09	−.01											
6. Triangle walk	−.01	−.02	−.08	.36	.32										
7. Arm swing	.08	.31	.04	−.07	.14	.07									
8. Arm circle	−.06	−.00	.13	−.01	.28	.20	.46								
9. Horizontal lines	−.11	.34	.09	−.27	.22	−.23	.12	.09							
10. Balance stick	−.14	−.27	.08	.24	.27	.16	.15	.06	−.14						
11. Finger spread	.11	−.11	−.06	.05	−.09	.24	.04	.05	−.17	−.01					
12. Ball balance	.17	.14	−.15	.05	.12	.06	.32	−.01	−.13	.15	.02				
13. Body sway	−.13	−.01	.12	−.02	.37	.22	.26	.27	.09	.23	.14	−.09			
14. Arm raising	.23	.04	−.19	−.15	.04	−.07	.22	.12	−.05	−.04	.08	.06	−.01		
15. Lower leg flexion	−.01	.03	−.08	.13	−.11	.05	.14	−.12	−.02	.19	.12	.15	.11	−.17	
16. Leg raising	.04	.10	.06	.11	−.06	.17	−.25	−.14	.09	−.24	−.07	.04	−.26	.09	.06

REFERENCES

1. Bass, Ruth I. An analysis of the components of tests of semicircular canal functions and of static and dynamic balance, *Research Quarterly*, **10**: 33, 1939.
2. Best, C. H., and N. B. Taylor. The Physiological Basis of Medical Practice. Baltimore: The Williams and Wilkins Co., 1945.
3. Espenschade, Anna, Robert R. Dable, and Robert Schoendube. Dynamic balance in adolescent boys, *Research Quarterly*, **24**: 270, 1953.
4. Fisher, Rosemary. A Study of *Kinesthesis in Selected Motor Movements.* Unpublished Master's Thesis, State University of Iowa, 1945.
5. Garrison, Levon. *An Experiment in Improving Balance Ability Through Teaching Selected Exercises.* Unpublished Master's thesis, State University of Iowa, 1953.
6. Honzik, C. H. Role of kinesthesis in maze learning, *Science*, **84**: 373, 1936.
7. Henry, Franklin M. Dynamic kinesthetic perception and adjustment, *Research Quarterly*, **24**: 176, 1953.
8. Lafuze, Marian J. Learning of fundamental skills by freshmen women of low motor ability, *Research Quarterly*, **22**: 149, 1951.
9. Lindsay, Doreen. *Relationship Between Measures of Kinesthesis and Learning of a Motor Skill.* Unpublished Master's thesis, University of California, Berkeley, 1952.
10. Meday, Helen Willcox. *The Influence of Practice on Kinesthetic Discrimination.* Unpublished Master's thesis, University of California, Berkeley, 1952.
11. Mullins, Peter M. *An Experimental Study of Kinesthesis in Learning in Basketball Throwing.* Master's thesis in process, State College of Washington, Pullman, 1954.
12. Mumby, H. Hugh. Kinesthetic acuity and balance related to wrestling ability, *Research Quarterly* **24**: 327, 1953.
13. Norrie, Mary Louise. *Relationship Between Measures of Kinesthesis and Motor Performance.* Unpublished Master's thesis, University of California, Berkeley, 1952.
14. Phillips, Bernath E. The relationship between certain phases of kinesthesis and performance during the early stages of acquiring two perceptuo-motor skills, *Research Quarterly*, **11**: 571, 1941.
15. Ragsdale, C. E. *How Children Learn the Motor Types of Activities.* 49th Yearbook, University of Chicago Press, Chicago, Illinois, 1950.
16. Rehling, Conrad H. *Teaching the Golf Swing by Using Blindfold Techniques.* Thesis in process, University of Florida, Gainesville, 1955.
17. Riley, Marie. *The Measurement of Balance in the Elementary Grades.* Unpublished Master's thesis, State University of Iowa, 1953.
18. Roloff, Louise L. Kinesthesis in relation to the learning of selected motor skills, *Research Quarterly* **24**: 210, 1953.
19. Scott, M. Gladys, and Esther French. *Evaluation in Physical Education.* St. Louis: C. V. Mosby Co., 1950.
20. Seashore, Harold G. The development of a beam-walking test and its use in measuring development of balance in children, *Research Quarterly*, **18**: 246, 1947.
21. Stevens, Mildred. *Measurement of Kinesthesis in College Women.* Unpublished doctoral dissertation, Indiana University, 1950.
22. Worchel, Philip. The role of the vestibular organs in space orientation. *J. Exp. Psych.* **44**: 4, 1952.
23. Wiebe, Vernon R. A study of tests of kinesthesis. *Research Quarterly*, **25**: 222, 1954.
24. Witte, Fae. *A Factorial Analysis of Measures of Kinesthesis.* Unpublished doctoral dissertation, Indiana University, 1953.
25. Young, Olive G. A study of kinesthesis in relation to selected movements. *Research Quarterly*, **16**: 277, 1945.

(Submitted 11/1/54)

Emotional Responses of Boys to Little League and Middle League Competitive Baseball[1]

ELVERA SKUBIC
University of California at Santa Barbara
Goleta, California

Abstract

This study was an experimental investigation of the emotional responses of boys 9 through 15 years of age to highly organized competitive baseball. The subjects were 75 boys who participated in Little League Baseball, 50 boys who were members of Middle League teams, and 80 boys who did not play on an organized team.

The Galvanic Skin Response test was employed to measure the emotional changes of boys participating in league baseball and to compare the results of these tests with those of the same boys participating in softball competition in physical education classes. Insofar as the Galvanic Skin Response test can be taken to be a valid measure of emotional excitation of boys of this age level, the results of this study suggest that youngsters were no more stimulated by competition in league games than they were by competition in physical education games.

THROUGHOUT THE NATION there is considerable interest in the promotion of competitive recreational activities for boys 15 years of age and younger and much has been said and written both pro and con in regard to the advisability of highly competitive activities for boys below the ninth grade.

Critics argue that there are harmful aspects of Little-League-type baseball, one of which is the emotional over-stimulation of boys which results from such factors as the presence of large numbers of spectators at games and undue emphasis on winning games. On the other hand, advocates of highly organized activities believe this type of experience to be more beneficial than harmful to youngsters.

Purpose of the Study

It appears that the controversy will continue to rage until evidence other than statements based on opinions is produced. It was the purpose of this study, therefore, to obtain by means of the Galvanic Skin Response test, scientific data regarding the emotional responses of boys in a variety of competitive situations. More specifically, it was the purpose of this experiment: 1. to test the emotional responses of boys before and after participation in Little and Middle League games and to compare these results with similar tests on the same boys before and after competition in physical edu-

[1]This study was made in partial fulfillment of the Ph.D. requirements at the University of Southern California, Los Angeles, California.

cation classes; 2. to compare the emotional response of league players with that of a group of non-players; 3. to determine whether there are differences in the emotional responses of boys at various ages; 4. to determine if there are differences in the emotional responses of boys to various intensities of competition; and 5. to determine the emotional reaction of boys to winning and losing games.

Terminology

Little League Baseball, Inc.—A national organization which promotes organized baseball for boys between the ages of 8 and 12.

Middle League Baseball, Inc.—A league which originated in Santa Maria, California, and promotes organized baseball for boys 12 through 15 years of age.

League Players—Subjects who were members of Santa Maria Little or Middle League teams.

Non-Players—Subjects who were not members of league teams either because they did not choose to belong or because they did not possess adequate skills to be chosen on teams. All boys who volunteered to take the tests in the schools where the experiment was conducted were included as subjects.

Galvanic Skin Response Test—A measure of changes in sweat gland activity. The G.S.R. affords an indication of the activity of the autonomic nervous system which is of major importance in the bodily changes associated with emotional states.

Emotional Response—The emotional change of a subject from one time to another time.

Emotional State—The emotional condition of a subject at one particular time.

Resting Level—A condition in which the subject is presumed to be under no particular stress.

Review of Literature

The desirability of highly organized athletics for boys below the ninth grade has become a major controversial issue. Although the elementary schools have for many years felt pressure to adopt the characteristics of the interscholastic program, the problem has just recently been brought into focus by the fabulous spread of Little League Baseball throughout the nation. The program is popular, and in 1954, over 70,000 boys participated on Little League teams. The pros and cons of the effects of such competition on boys has been discussed, and a search of the literature shows that a lack of agreement exists between educators, coaches, child specialists, and sportswriters in regard to policies and practices, advantages and disadvantages of such participation.

Those who favor Little League activities such as Ridinger (36), Hale (13), Zorn (34), Aromandi (4), and Thees (14), believe that the movement was born out of the needs and desires of young boys and that there are

tremendous possibilities for youngsters to develop personal security and emotional stability. Questionnaire studies done by the National Recreation Association (32), the Fresno Recreation Department (15), and Phebe Scott (37) indicate that the majority of parents whose boys play league baseball and many recreation departments throughout the country approve Little League Baseball for boys.

On the other hand, there are those who feel that intense competition is not only educationally unsound, but also detrimental to the physical, psychological, and emotional well-being of young boys. Lowman (24, 25), McNeely (26), Morris (29), Keene (23), Morehouse and Miller (28) are a few who cite evidence to indicate that boys may be harmed physically by such activities. Jersild (16), Fait (10), Shaffer (38), and Solomon (39) have written of the psychological effects which may result from improperly conducted competition, and Montgomery (27), Solomon (39), and Bucher (5) are a few who have written of emotional harm from intense competition.

Many groups, through their national organizations, have passed resolutions against highly organized programs. The North Carolina Recreation Commission (35), the National Committee on Cooperation in Health Education (30), the National Federation of High School Athletic Associations (31), the National Recreation Association (33), the American Association for Health, Physical Education, and Recreation (1), and the National Conference on Program Planning in Games and Sports for Boys and Girls of Elementary School Age (2) represent many members who believe a broad, well-conducted intramural-type program would better serve the needs of boys than an interscholastic-type program.

The literature revealed very few experimental studies that attempted to determine the effects of competition on emotion. Johnson (19) used a psychogalvanometer in combination with word association tests in order to study pre-competitive emotional behavior of college athletes and found that a pre-contest stress situation will provoke greater emotional responses to critical words in sports than the undisturbed condition. In another study, Johnson (18) reported that many sports champions report excruciating emotional distress before a contest even though they appear calm.

The analysis of the literature strongly suggests the need for objective evidence on all phases of competition for youngsters. Since no studies were found which tested the emotions of boys in the 9-to-15-year age group, this investigation was an attempt to meet this need by studying one phase of the problem—the immediate emotional responses of young boys to competition.

Limitations of the Study

When an experiment is conducted in the midst of real life situations rather than in the laboratory situation, precise controls are not always possible. In the present study, most of the testing took place in school buildings and on the play fields. It was, of course, impossible to eliminate the normal noises, the ringing of bells, and the interruptions that are part of the school or play-

ground situation. Also, it was difficult to control such factors as the subject's time of eating, the amount of exercise taken prior to the tests, and the possibility of extraneous emotional stimuli. However, a large number of subjects were involved and the small errors caused by lack of precise controls were probably minimized.

Another serious problem involved the study of emotion itself. Emotions are extremely complex and neither the precise function nor the exact nature of the emotion is known. In studies such as the present one, emotion must be dealt with indirectly by way of physiological indicators which are known to vary more or less systematically with certain psychological states. Plainly, research findings based upon the physiology of emotion must be interpreted with care.

Research Procedures

The Galvanic Skin Response Test. When the emotional state of an individual changes, his body reflects this change in many ways. The most obvious indicators are the respiratory rate, muscular tension, pulse rate, and blood pressure. However, many other bodily changes are associated with emotional states and one of the most significant is the resistance of the skin to the flow of electric current. During some emotional states, the sweat glands of the palms become active and offer greater resistance to a current passed through the skin. This resistance can be measured and is called the galvanic skin response (12). It should be noted that neither galvanic skin response nor any other physiological indicator measures emotion *per se.* Some physiological states can be measured to ascertain the extent of certain bodily changes but these are merely suggestions of concomitant emotional excitation.

In the present study, the instrument used was a sensitive galvanometer-type instrument. Attached to this were zinc-plated electrodes of the type used by the medical profession. Microhm jelly was used to insure contact between the electrode and the skin.

The Subjects. The subjects were 206 boys ranging in age from 9 through 15 who lived in Santa Maria, California, and adjacent towns. Of this total, 75 boys were 9 to 12 years of age and were members of Little League teams during the summer of 1953. Fifty-one boys, ranging in age from 12 through 15, played on Middle League teams. A group of 80 non-players, 9 through 15 years of age, were chosen at random and used as a control group.[2]

Administration of the Tests. Each subject to be tested was seated in a chair with his arms resting comfortably on a table. The palm of the right hand and a spot on the lower left arm one inch below the crease of the elbow were cleansed with alcohol and rubbed briskly with electrode jelly. The electrodes were applied to the palm and the arm and the reading recorded in ohms.

[2]For their wholehearted co-operation in facilitating the collection of data, special gratitude is extended to William Ellis, founder of Middle League, Inc., the coaches of Little League and Middle League teams, the principals and teachers of the Santa Maria schools, and all the boys who participated in the study.

The following tests were administered:

1. All subjects were tested immediately before softball competition in physical education classes, immediately after their class and 1½ hours after class.

2. The subjects who played on Little or Middle League teams were tested immediately before competition in baseball, immediately after the game was over, and 1½ hours after the game. This procedure was repeated at three of their games.

3. All subjects, players and non-players, were tested while in the most relaxed state possible and while under no known stress. Fifty of these subjects were retested in approximately ten minutes in order to check the reliability of the test.

Analysis of the Data

Reliability. The correlation of the two resting level galvanic skin response scores for 50 subjects was .80 with a p.e. of .03. This finding is consistent with the high reliabilities found by other workers such as Elbel and Ronkin (10), Jones (23), Grings (13), Croslans and Beck (8), and Johnson (20).

Resting Level Scores. The range of galvanic skin response scores for all subjects while at rest was 6,700 to 140,000 ohms. The population mean was 17,400 ohms with a standard deviation of 10,700 ohms. Analysis of variance indicated that there was variability in scores from one individual to another.

Table 1 shows the means for each age group. Table 2 shows a significant difference in scores between age groups at the 5 per cent level of confidence. The average resistance of individuals at different ages seems to be different.

TABLE 1

Resting Level Means for Each Age

Age	No. Cases	Mean in Ohms	Standard Deviation
9	19	1.60	.67
10	37	1.46	.83
11	38	1.59	.76
12	31	1.73	.91
13	29	2.73	3.06
14	30	1.65	.67
15	26	1.52	.58
Total	210	1.74	1.07

TABLE 2

Analysis of Variance for Resting Level of All Subjects

Source of Variance	Sum of Squares	D. F.	Var. Est.	F	P
Bet. age groups	33.2	6	5.53	3.30	.05
Within groups	341.65	203	1.68	—	—
Total	374.85	209			

In general, the youngest boys tended to show the greatest skin conductivity and this decreased with age. The 13-year-old boys showed the least emotional response and, although the reason for this is not clear, it is possible that this phenomenon is in some way related to maturation and puberty.

The resting scores of those boys who played on league teams were compared to scores of non-players. In general, the players at all ages evidenced greater skin resistance (less change from the resting level) than the non-players at all ages. See Table 3.

TABLE 3

Differences Between Players and Non-Players for Resting Level

Source of Variance	No. Cases	Mean	S. D.	t	P
Age 9					
Players	9	1.73	.18	+.2089	—
Non-players	9	1.64	.22		
Age 10					
Players	21	1.74	.34	+6.86	.01
Non-players	17	1.35	.035		
Age 11					
Players	28	1.77	.15	+3.24	.01
Non-players	11	1.13	.13		
Age 12					
Players	23	1.98	.24	+3.71	.01
Non-players	8	1.00	.081		
Age 13					
Players	16	2.72	.58	−.02	—
Non-players	13	2.75	1.10		
Age 14					
Players	19	1.97	.13	+4.20	.01
Non-players	11	1.12	.15		
Age 15					
Players	16	1.88	.12	+6.22	.01
Non-players	10	.96	.08		

G.S.R. Scores Relating to Physical Education. Analysis of variance was computed for Little and Middle League players for the three physical education tests (immediately before and after class and 1½ hours after class), and it was found that significant differences at the 1-per-cent level of confidence existed among the three tests.

For the non-players, there were significant differences between the test for boys aged 9, 10, and 11 but not for those 12 through 15.

In comparing the players with the non-players on the three tests, the Fisher t scores showed in general that players tend to show less emotionality than the non-players. See Table 4.

G.S.R. Scores Relating to League Games. Seventy-five subjects who played on teams were tested three times on the day of a league game (immediately before and after the game and 1½ hours after the game). Analysis of variance showed that for almost every age there were large differences in

TABLE 4

Difference Between Players and Non-Players on Three Physical Education Tests

Test	Age	T	Probability
Test 1—before physical education	9	4.10	.01 Players more relaxed
	10	.08	—
	11	.37	—
	12	.34	—
	13	.75	—
	14	1.30	—
	15	.008	—
Test 2—After physical education	9	3.39	.01 Players more relaxed
	10	.62	—
	11	.72	—
	12	1.72	—
	13	.15	—
	14	2.07	.05 Players more relaxed
	15	.024	—
Test 3—1½ hours after physical education	9	.97	—
	10	.067	—
	11	3.08	.01 Players more relaxed
	12	1.86	.01 Players more relaxed
	13	.35	—
	14	2.72	.01 Players more relaxed
	15	3.83	.01 Players more relaxed

scores between the three tests and significant differences between the scores of individual players. See Table 5.

Comparison of Physical Education Scores with League Scores. There appeared to be a general pattern of emotional response for the four tests taken on one day either for physical education competition or league competition. The resting level tests, as would be expected, showed the most relaxed state of the players and non-players. Just before participation in a game, league players were in an excited state which generally increased right after the game was over. On the third test 1½ hours after the completion of the game, the players consistently reached a higher degree of relaxation than they had before the physical education and league games. However, this level of relaxation was still far below the resting level scores.

With the exception of the nine-year-olds, the non-players were considerably more excited 1½ hours after their physical education games than they were before the start of the games.

When emotional reactions of subjects before a physical education game were compared with those for league games, no significant differences were found. The players showed no greater skin response in anticipation of a league game than they did before competition in physical education classes.

Of the tests taken immediately after the games, the 10-, 11-, and 13-year-old players showed more skin response after physical education competition

TABLE 5

Analysis of Variance for the Game Situation for Players (Average of Three Games)

Source of Variance	Mean	Sum of Squares	D. F.	Var. est.	F	P
Age 9 (9 subjects)						
Between tests	1.50	1.04	3	.34	.40	—
Between players		4.15	32	.51	.60	—
Residual		2.05	24	.85		
Age 10 (21 subjects)						
Between tests	1.88	2.05	3	.68	2.21	—
Between players		27.86	20	1.39	4.52	.01
Residual		19.37	63	.31		
Age 11 (28 subjects)						
Between tests	1.51	4.70	3	1.56	3.54	.05
Between players		37.53	109	1.32	3.00	.01
Residual		35.70	81	.44		
Age 12 (23 subjects)						
Between tests	1.58	7.22	3	2.40	5.88	.01
Between players		26.29	91	1.19	2.91	.01
Residual		26.98	66	.41		
Age 13 (16 subjects)						
Between tests	1.64	26.07	3	8.69	7.12	.01
Between players		30.64	15	2.04	1.67	—
Residual		55.34	45	1.22		
Age 14 (19 subjects)						
Between tests	1.36	9.52	3	3.17	41.16	.01
Between players		7.00	18	.38	5.03	.01
Residual		4.18	54	.08		
Age 15 (16 subjects)						
Between tests	1.40	6.64	3	2.21	11.63	.01
Between players		7.52	15	.50	2.63	.01
Residual		8.81	45	.19		

than they did after a league game. Also, the 10-, 11-, and 12-year-old players showed more skin response 1½ hours after their league games. At all other ages, there were no significant differences. See Table 6.

Emotional Reactions to Winning and Losing Games. Galvanic skin response scores of players taken immediately after losing league games were compared to the scores taken immediately after winning games, and the results as shown in Table 7 indicated that players show greater skin response after winning than after losing a game.

Comparison of G.S.R. Scores for a League Game, a Physical Education Game and a Championship Game. There appears to be little difference in the galvanic skin response of boys to different intensities of competition. The only significant difference, at the 1-per cent level of confidence, was the greater excitability of Little League players before a semi-final state championship game as compared to a quarter final game.

TABLE 6

Analysis of Variance for Players in League Games and Physical Education Games

Source of Variance	t	p
Age 9		
Before a physical education game—before a league game	.11	—
After a physical education game—after a league game	1.36	—
1½ hours after a phys. educ. game—1½ hours after a league game	.74	—
Age 10		
Before a physical education game—before a league game	.47	—
After a physical education game—after a league game	1.90	+.05
1½ hours after a phys. educ. game—1½ hours after a league game	2.50	+.05
Age 11		
Before a physical education game—before a league game	1.49	—
After a physical education game—after a league game	2.74	+.01
1½ hours after a phys. educ. game—1½ hours after a league game	1.84	+.05
Age 12		
Before a physical education game—before a league game	1.49	—
After a physical education game—after a league game	1.51	—
1½ hours after a phys. educ. game—1½ hours after a league game	2.83	+.01
Age 13		
Before a physical education game—before a league game	1.03	—
After a physical education game—after a league game	2.33	+.05
1½ hours after a phys. educ. game—1½ hours after a league game	.39	—
Age 14		
Before a physical education game—before a league game	.27	—
After a physical education game—after a league game	.51	—
1½ hours after a phys. educ. game—1½ hours after a league game	.47	—
Age 15		
Before a physical education game—before a league game	1.59	—
After a physical education game—after a league game	.08	—
1½ hours after a phys. educ. game—1½ hours after a league game	1.22	—

TABLE 7

Significance of the Difference Between Winning and Losing a League Game

Immediately After the Game			1½ Hours After the Game		
No. Cases	t	p	No. Cases	t	p
132	5.53	.01	132	2.25	.05

Summary and Conclusions

The present experiment was designed to collect objective data regarding the effects of competition on the emotions of boys between the ages of 9 and 15. The galvanic skin response test was employed to measure emotional changes of boys who were participating in Little League and Middle League Baseball and to compare the results of these tests with those of the same boys participating in softball competition in physical education classes.

Many critics of Little League Baseball have stated that highly organized competition places physical and emotional strains on boys which may be harmful to their general welfare. Insofar as the G.S.R. can be taken to be a

valid measure of the emotional excitation of boys of this age level, the results of the present study suggest that youngsters were no more stimulated by competition in league games than they were by competition in physical education games.

It should be remembered that this study was concerned with only one phase of the total problem of competition—the immediate effects of competition on emotionality. In order to completely solve the problem of highly organized competition, data must be gathered relative to the physical, sociological, psychological, and economic aspects of competition. Furthermore, to resolve the specific controversy concerning the emotional effects of competition, it is necessary that additional data be secured, particularly in regard to the influence of emotion on personality now and later in life.

Within the limits of this study of boys 9 through 15 years of age who played Little League and Middle League competitive baseball in Santa Maria, California, and within the limits of information about emotion as derived from the galvanic skin response test, the evidence presented supports the conclusion that Little-League-type competition is no more stimulating than competition in school physical education.

REFERENCES

1. American Association for Health, Physical Education, and Recreation. *Desirable Athletic Competition for Children.* A report of the joint committee on athletic competition for children of elementary and junior high school age. Wash., D. C.: the Association, 1952.
2. American Association for Health, Physical Education, and Recreation. *Summary of the National Conference on Program Planning in Games and Sports for Boys and Girls of Elementary School Age, May 25-26, 1953.* Wash., D. C.: the Association, 1953.
3. Arnold, Magda B. Physiological Differentiation of Emotional States, *Psychological Review,* **152:** 35-48 (1945).
4. Aromandi, Nick. Boro Little Leaguers Hit by Attack Made by Educator, *Bronx Press Review,* New York, (Aug. 6, 1953).
5. Bucher, Charles A. Little League Can Hurt Your Boy, *Look* Magazine, **18:** 3 (Sept. 22, 1953).
6. Butts, Bevier. Little League—Yes or No? *Illinois Recreation Association Bulletin,* No. 31; Waukegan, Ill. (1953).
7. Castle, Ames. What Oakland Parents Say About Little League Ball, *Industrial Sports Journal,* **13:** 13 (May 15, 1952).
8. Crosland, Harold, and Lester Beck. Objective Measurements of Emotion, *Quarterly of the Univ. of Oregon,* Eugene, Oregon: *Psychological Series,* **1:** 3 (Dec. 1931).
9. Elbel, E. R., and R. R. Ronkin. Palmar Skin Resistance as a Measure of Physical Fitness, *American Jl. of Physiology,* **147:** 1-12 (Sept. 1946).
10. Fait, Hollis. *An Analystical Study of the Effects of Competitive Athletics Upon Junior High School Boys,* Ph.D. dissertation, University of Iowa (Aug. 1951).
11. Gish, Robert L. Electropsychometer, *Electronics,* **11:**: 58-61 (May 1953).
12. Grings, W. M. *Methodological Consideration Underlying Electrodermal Measurements,* Dept. of Psychology, University of Southern California, Los Angeles, Calif.
13. Hale, Creighton. Athletic Competition for Children? *Athletic Journal,* **34:** 190-192 (Jan. 1954).

14. Herald Tribune. New York, N. Y., (Sept. 6, 1953).
15. Holman, Edward. *A Study of Little League Baseball in Operation—1951.* Recreation Director, City Hall, Fresno, Calif. (1952).
16. Jersild, A. T., and H. E. Jones. Adolescent Growth Study, *Jl. of Consulting Psychology*, **3**: 157 159; 177-180 (1939).
17. Johnson, Warren. Emotional Upset in Athletes, *Athletic Journal*, **32**: 7-10 (Nov. 1951).
18. ———. Psychogalvanic and Word Association Studies of Athletes, *Research Quarterly*, **22**: 427-433 (Dec. 1951).
19. ———. *A Study of the Emotion of College Athletes*, Ph.D. dissertation, Boston University (1950).
20. ———. A Study of Emotion Revealed in Two Types of Athletic Contests, *Research Quarterly*, **20**: 72-79 (1949).
21. Johnson, Warren, and J. Harmon. Emotional Reaction of College Athletes, *Research Quarterly*, **23**: 391-397 (Dec. 1952).
22. Jones, H. E. Galvanic Response in Infancy, *Child Development*, **1**: 106-110 (1930).
23. Keene, C. H. Sports and Games in the School Schedule, *Hygeia*, **20**: 557-558 (July 1942).
24. Lowman, C. L. A Consideration of Teen-Age Athletics, *Journal of the American Association for Health, Physical Education, and Recreation*, **12**: 398-399 (Sept. 1941).
25. ———. The Vulnerable Age, *Journal of the American Association for Health, Physical Education, and Recreation*, **18**: 635-636 (Nov. 1947).
26. McNeely, Simon. Of "Mouse" and Men, *Journal of the American Association for Health, Physical Education, and Recreation*, **23**: 27-28 (Dec. 1952).
27. Montgomery, Katherine. *A Giant Controversy Over Midget Sports*, Director of Physical Education, Florida State College, Tallahassee, Florida.
28. Morehouse, L. E., and A. T. Miller. *Physiology of Exercise*, Philadelphia: C. V. Mosby Co., 1948.
29. Morris, Morty. Don't Rush Your Kids, *Journal of the American Association for Health, Physical Education, and Recreation*, **23**: 10 (Oct. 1952).
30. National Committee on Co-operation in Health Education. *Suggested School Health Policies.* Chicago, Ill.: American Medical Association, 1945.
31. National Federation of High School Athletics. Cardinal Athletic Principles—A Policy Statement. *Journal of the American Association for Health, Physical Education, and Recreation*, Sept. 1947.
32. National Recreation Association. A Survey of Recreation Executives on Organized Competition in Sports and Athletics for Boys Under Twelve.
33. ———, Standards Committee. Competitive Athletics for Boys Under Twelve. *Recreation*, **45**: 489-491 (Feb. 1952).
34. News Palladium, Benton Harbor, Michigan (Aug. 4, 1953).
35. North Carolina Recreation Commission. *Competitive Athletics for Children*—Policy Statement. Education Building, Raleigh, N. C. May 1952.
36. Ridinger, Bill. Midget Sports, *Industrial Sports Journal*, **12**: 40-41 (Dec. 1951).
37. Scott, Phebe, Attitudes Toward Athletic Competition in Elementary Schools. *Research Quarterly*, **24**: 352-361 (Oct. 1953).
38. Shafer, Roy. A Study of Thought Processes in a Word Association Test, *Character and Personality*, **13**: 212-227 (1945).
39. Solomon, Ben. Little League—Menace or Blessing? *Youth Leaders Digest*, **15**: 161-213 (Feb. 1953).

(Submitted 8/13/54)

Game Results as a Criterion for Validating Basketball Skill Test

FRANCIS STROUP
Southern State College
Magnolia, Arkansas

Abstract

In this study the scores of the competing teams in 31 ten-minute basketball games are compared with the skill score averages of the teams. The skill score averages were computed from scores on passing, dribbling and shooting tests. Approximately 84 per cent of the games were won by the team with the higher skill score average.

THE SELECTION of a suitable criterion for establishing test validity is a major problem in test construction. Attempts to express numerical coefficients of validity have stimulated the use of a variety of such criteria. The criteria used in validating physical education performance tests have been predominantly judgment ratings made by qualified observers (1, 5, 6, 7). However, other criteria have been used.

Dyer used relative final positions of subjects in a round robin tournament as the criterion for establishing the validity of a tennis test (4). Criteria based on game results have also been used by Lockhart and McPherson (9) in badminton, by Miller (11) in badminton, by Lockhart and Mott (10) in table tennis and by Cornish (3) in handball.

But the use of game results as a criterion in validating tests has been limited largely to individual and dual sports. This limitation is attributable to the relative ease of controlling factors in games involving one or two performers.

The Problem

Two related purposes provided the problem for investigation in this study. They were: (1) to demonstrate the use of a validation technique for a team sport test in which game results were used as the criterion, and (2) to establish the validity of an administratively economical test for equating teams.

As a means of improved validation. The demonstration of a technique in which game results are used as a criterion for validation would be an important step in attaining greater confidence in sport skill tests. Regardless of the magnitude of the derived coefficient of correlation between scores on a test and a selected validity criterion, the test will not inspire confidence if there is doubt regarding the appropriateness of the criterion. And confidence is necessary if a test is to receive wide use.

No criterion for determining the validity of a sport test seems more appropriate than the results of contests in that sport. If it is possible to show that such a criterion can be suitably employed, tests validated by the use of such a criterion should inspire greater confidence and consequently should receive wider use.

Demonstration of the use of such a criterion could perform a second function. If game results are used successfully as a criterion for validating a basketball test, the demonstration of that use should suggest ways of using a similar criterion for validating tests in other team sports. Thus, further research and consequent improvement in physical education sports testing might be stimulated.

As a means of equating teams. An easily administered test for equating basketball teams has many potential uses. Most of the benefits claimed for sports participation are lost unless a high level of interest by the participant is maintained. Grossly unequal competition is a great deterrent to such interest. A simple test for reducing the inequality of teams would be useful, not only in school classes and intramural programs, but also in other institutions such as vacation camps and military installations where participants are grouped for competition on short acquaintance.

The achievement of these two purposes seemed sufficiently important to justify this study.

Procedure

In order to achieve these purposes, a study was made using men enrolled in four sections of Freshman and Sophomore physical education classes at Southern State College as subjects.

The test items. Since the use of an administratively economical test was one of the purposes of this study, it was decided to use a three-item test. Most investigators of basketball skill tests have recommended a battery of five items or less (7, 8, 12). Glassow, Colvin, and Schwarz indicated that the fourth and fifth items investigated by them made no contribution to the validity of the battery (6). The three items selected for use in this study were goal shooting, wall passing, and dribbling.

In the goal-shooting test, the subject was allowed to stand as near as he wished to the basket and shoot as many baskets as possible in one minute, retrieving the ball each time himself. No penalties were invoked for rules violations since the time limit penalized and discouraged such infractions.

In the wall-passing test, the subject stood behind a line six feet from a wall and passed against the wall as many times as possible in one minute. It was considered a miss to bat the ball instead of catch it or to move beyond the restraining line while handling the ball. The score was the number of legal passes made in one minute.

The dribbling test required the subject to dribble alternately to the left and right of bottles placed in line and 15-ft. apart on a 90-ft. court, circle the end bottle, and continue in this manner for one minute. It was considered a miss to knock over a bottle or to not pass a bottle on the proper side. The score was the number of bottles properly passed in one minute.

Acceptable coefficients of reliability and validity have been reported for similar or identical test items by other investigators (7, 8, 12, 6). The criterion for validity in these studies was either subjective ratings or scores on a larger battery of tests.

Scoring procedures. Previous experimentation with the test items used in this study has led to the development of scale scores with letter equivalents for performances on each item. These scale scores were based on a theoretical division of the population of the normal curve. The scale scores, letter equivalents and division of the population are indicated in Table 1. A subject's raw scores for the three items were converted to scale scores which were then averaged to obtain his basketball *skill score.* A table used for this conversion is shown in the Appendix, page 357.

TABLE 1

Corresponding Scale Scores, Letter Grades and Per Cent of Population of Normal Distribution

Scale Score	Letter	Per Cent
91-100	A	7
81- 90	B	24
71- 80	C	38
61- 70	D	24
	F	7

As a part of this study, ten-minute basketball games were played in four team sport sections of physical education classes for men. These games were spaced over the last six weeks of the semester and team members for each game were selected at random. A scorekeeper kept the lineups and scores.

At the close of the semester, the three-item basketball test was administered to class members. Skill scores for the subjects were computed as previously described and each subject's skill score was inserted before his name on the score sheet of each game in which he participated.

Skill scores for the members of each team were averaged and comparisons of competing teams were based on *average skill scores* and *game scores.*

Results

Forty-one ten-minute games were played during the study. The three-item basketball test was administered to 121 subjects. Subjects' skill scores derived from test performances ranged from 52 to 84 with a mean of 72.65.

Owing to drop-outs and injuries, all men participating in the games did not take the test. Skill scores were available for all members of 75 teams and averages for these teams ranged from 65.2 to 78.6 with a mean of 73.31.

There were 38 games for which average skill scores were available for both competing teams. In one of these, the teams had identical skill score averages and played to a tie game. In six games, the competing teams had equal average skill scores or equal game scores. There was a difference in both skill scores and game scores for the competing teams in 31 games.

The average skill score differences for these thirty-one games ranged from 0.4 to 10.2 with a mean of 3.45. The game score differences ranged from 2 to 18, with a mean difference of 7.36.

Of the 31 games in which differences both between average skill scores and between game scores for the competing teams were observed, 26 (83.87%) of the games were won by the team having the higher average skill score. The five games with outcomes counter to this pattern had average skill score differences of 0.8, 1.0, 2.2, 4.6, 6.6 and respective game score differences of 6, 6, 4, 4, and 6. In no case did a team having an average skill score advantage of more than 6.6 lose a game.

Summary

This study was designed: 1. to demonstrate the use of a validation technique for a team sport test in which game results were used as the criterion, and 2. to establish the validity of an administratively economical test suitable for equating basketball teams.

To achieve these objectives, members of four physical education classes played ten-minute games during basketball participation periods. At the close of the semester a basketball test consisting of passing, shooting, and dribbling was administered to them.

Raw scores on the test items were converted to scale scores which were then averaged to find each subject's basketball skill score. This skill score was inserted before each subject's name on the score sheet of games in which he had participated. By this method, average skill scores of competing teams could be compared with the scores resulting from the games.

For instances in which there were differences both in average skill scores and in game scores for the two competing teams, 83.87 per cent of the games were won by the team having the higher skill score. In cases not complying with this pattern the differences between average skill scores and the differences between game scores were relatively small.

Conclusions

From the results of the study, these conclusions were drawn:

1. The procedure employed in this study demonstrates the use of ten-minute game results as a criterion for validating a team sport test.

2. Average skill scores derived from scores on the three-item test are a valid measure of team strength in basketball because relative skill scores of competing teams was related to the ability to win ten-minute games.

3. The test appears to be a practical instrument for equating teams, because knowledge of average skill score differences for competing teams in games not involving ties would have allowed correct prediction of the outcome of approximately 84 per cent of these games. In the remaining 16 per cent, differences in average skill scores and game scores were usually small.

REFERENCES

1. Brace, David K. *Measuring Motor Ability.* New York: A. S. Barnes and Co., 1927.
2. Cozens, F. W. *The Measurement of General Athletic Ability in College Men.* Eugene, Oregon: University of Oregon Press, 1929.
3. Cornish, Clayton. A study of measurement of ability in handball, *Research Quarterly,* **20**: 215-22, 1949.
4. Dyer, Joanna T. Revision of backboard test of tennis ability, *Research Quarterly,* **9**: 25-31, 1938.

5. EDGREN, H. D. An Experiment in the Testing of Ability and Progress in Basketball, *Research Quarterly*, **3**: 159-71, 1932.
6. GLASSOW, RUTH B., VALERIE COLVIN and MARGUERITE SCHWARZ. Studies in measuring basketball playing ability of college women, *Research Quarterly*, **9**: 60-68, 1938.
7. JOHNSON, L. W., *Objective Basketball Tests for High School Boys.* Unpublished Master's thesis, State University of Iowa, Iowa City, Iowa, 1934.
8. LEHSTEN, NELSON. A measure of basketball skills in high school boys, *The Physical Educator*, **5**: 103-9, 1948.
9. LOCKHART, AILEENE, and FRANCES A. MCPHERSON. Development of a test of badminton playing ability, *Research Quarterly*, **20**: 402-5, 1949.
10. LOCKHART, AILEENE, and JANE A. MOTT. Table tennis backboard test, *J. Health and Physical Education*, **17**: 550-2 (Nov. 1946).
11. MILLER, FRANCES A. A badminton wall volley test, *Research Quarterly*, **22**: 208-13, 1951.
12. YOUNG, GENEVIEVE, and HELEN MOSER. A short battery of tests to measure playing ability in women's basketball, *Research Quarterly*, **5**: 3-23, 1934.

APPENDIX

Scale Score Equivalents for Performances in Shooting, Passing and Dribbling

Shooting	Passing	Dribbling	Scale score	Shooting	Passing	Dribbling	Scale score
6	53	27	51	24	78	42	76
7	55		52				77
8	56	28	53	25	79	43	78
9	57	29	54	26	80		79
	59	30	55	27	81	44	80
10	60	31	56		82		81
11	61		57	28		45	82
12	62	32	58	29	83		83
13	64	33	59		84	46	84
14	65	34	60	30	85		85
	66		61		86	47	86
15		35	62	31	87		87
16	67		63	32	88	48	88
	68	36	64		89	49	89
17	69		65	33	90	50	90
	70	37	66	34	91		91
18			67	35	93	51	92
19	71	38	68	36	94		93
	72		69	37	95	52	94
20	73	39	70		97		95
21			71	38	98	53	96
	74	40	72	39	99		97
22	75		73	40	100	54	98
23	76	41	74	41	102	55	99
	77		75	42	103	56	100

(Submitted 10/6/54)

Emotion in Fitness Tests

PHILIP V. WELLS
Prudential Insurance Company of America
Newark, New Jersey

Abstract

The paper presents a method of estimating the emotional factor in flarimeter tests on athletes, by analyzing an index of responses to standard stimulus.

FITNESS TESTS have failed to meet expectations. In studying flarimeter tests[1] on athletes it has become evident that the real reason for this is that such tests are disturbed by emotion. This paper presents a method of estimating the emotional factor.

Standard Stimulus

Standard stimulus for athletes (4) consisted, first, of Master's standard step exercise for 90 seconds. The subject was then seated with dispatch, began a flarimeter blow 10 seconds after exercise which was stopped after blowing 30 seconds. The subject then sat still until the final blow was begun at 2½ minutes after exercise, which was held until the systolic blood pressure rose 30 mm.

Systolic blood pressure was read at 1 minute after exercise. Let S_1 represent its value in mm of mercury. Diastolic blood pressure was taken immediately after the systolic. Call this D_1. Similarly, S_2 and D_2 were the readings at 2 minutes after exercise, and R′ the pulse rate per minute, from a 15-sec. count begun at 1½ minutes after exercise.

D_2 was usually equal to the diastolic at rest before exercise, indeed D_1 was seldom appreciably higher, and the pulse rate R′ was usually slightly below its pre-exercise value. Most of the response was in S_1. It was soon discovered that the reason for this was the powerful vagus action of holding the breath in the first blow after exercise. Continuous electrocardiographic records of Lead II showed sudden vagus slowing whenever the heart accelerated enough, most often toward the end of the blow but sometimes early in the blow or after it stopped.

Flow Index

To avoid bias, the responses to standard stimulus were studied separately for over 15 years, in spite of the obvious tendency of the systolic pressure to vary reciprocally with the pulse rate. Finally, the five measures (S_1, D_1, R′, S_2 and D_2) were combined into an index, $F' = P'R'/100$, where $P' = (P_1 + P_2)/2$ and $P_1 = S_1 - D_1$, $P_2 = S_2 - D_2$. This seemed the least empirical formula to use. It was not intended as a law of physiology, al-

[1]The flarimeter is a clinical instrument for testing circulatory fitness (2, 3). The subject blows through a standard orifice at a standard pressure. The examiner can follow the systolic blood pressure during the blow, to measure the circulatory response.

though the "pulse product" has been studied seriously as a measure of blood flow ever since its introduction by Erlanger and Hooker in 1904 (5). In view of this, it seems permissible to call $F' = P'R'/100$ a "flow index," in arbitrary units. Its median value for Princeton University athletes was about 30 after standard stimulus, while the median value of $F = PR/100$ sitting at rest before exercise was about 25, where $P = S - D$ was the pulse pressure and R the pulse rate.[2]

The flow index is much better than blood pressure or heart rate as a measure of the response to standard stimulus. Momentary changes in pulse pressure or rate are often erratic when their product, the flow index, is steady. To determine the relation of F′ to F, individual medians of 83 tests on 31 Princeton varsity swimmers at peaks of training (selected out of 226 tests on them in all conditions) were plotted in a scatter diagram, shown in Fig. I. A least-squares fit gave the regression line $F' = 1.2F$, with a correlation $r = 0.87$, but the fit of line $F' = F + 5$ is quite as good, as shown by

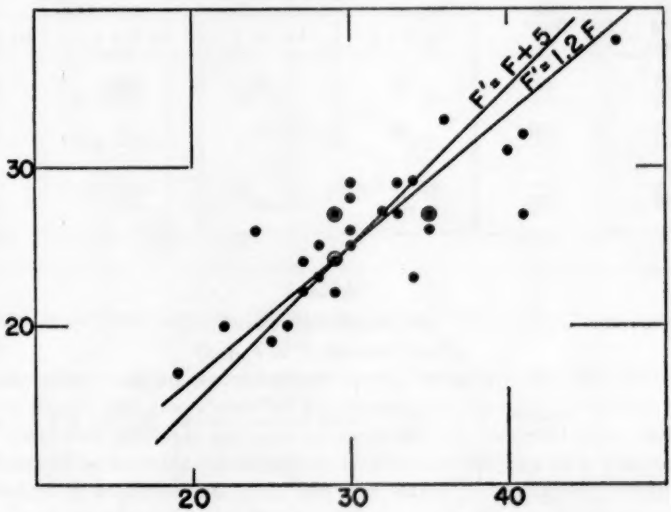


FIGURE I. Scatter Diagram of the Flow Index.
Abscissas — F, the index before exercise.
Ordinates — F′, the index after standard stimulus.
(Circles represent second points identical in values with dots inside)

[2]Since P′ is the average of P_1 and P_2, it approximates the pulse pressure simultaneous with the pulse rate R′. So the product F′ has some physiologic meaning. It is not merely a statistical index. Physiologists point out (1) that F′/F measures relative blood flow only when arterial elasticity does not change, but the fact that the diastolic blood pressure after standard stimulus is usually nearly the same as before exercise is presumptive evidence that F′ is a fairly good crude index of blood flow under the conditions of the test.

Table 1, and this line is more practical. A summary of its agreement with individual medians is shown in Table 2. These tables give the average differences between observed and calculated values of F'. Disregarding sign the average deviations are as shown in the columns headed (± a.d.).

Notice that seven of the 31 points in Fig. I actually fall on the line $F' = F + 5$. It is quite obvious that standard stimulus produces an almost constant increment of five units in the flow index. However, there is a tendency for the increment to increase with the index. Out of 204 individual medians, 92 with F' less than 30 give a median increment of three, while

TABLE 1

Comparison of Fit of Formulas
$F' = F + 5$ and $F' = 1.2\ F$
to Individual Medians on 80 Varsity Swimmers

No. Men	No. Tests	Av. (Obs.-Calc.) F'		± a.d.	
		By F + 5	By 1.2 F	By F + 5	By 1.2 F
31	226	.29	.19	2.6	2.7
49	170	.10	—.57	4.0	4.4
80	396	.16	—.28	3.4	3.7

TABLE 2

190 Individual Medians
(Fit of Formula $F' = F + 5$)

No. Men	Varsity Athletes	No. Tests	Medians		Av. Dev. F' (Obs.-Calc.)	a.d. ±
			F	F'		
31	Swimmers	226	26	30	.29	2.6
49	"	170	27	32	.10	4.0
80	"	396	26	30	.16	3.4
88	Track	348	24	29	—.04	3.5
10	Bk 1937	73	24	30	.9	3.1
12	Crew 1936	81	31	36	.75	1.9
190		898			.20	3.3

54 with F' 35 or over give a median increment of eight units. The same excitability which produces above-average flow index also increases the response to standard stimulus. The expected increment in the index required to clean up the oxygen debt, therefore, is about five units or 20 per cent.

To make sure there was no bias in the medians, the formula was applied to unselected single tests. The 200 latest tests on swimmers, the 100 latest on runners, all the tests on the 1937 basketball team and two varsity crews were taken, with results shown in Table 3. The fact that the formula fits

TABLE 3

Fit of Formula F' = F +5 to 518 Unselected Tests

No. Men	Varsity Athletes	No. Tests	Medians		Av. Dev. F' (Obs.-Calc.)	a.d. ±
			F	F'		
92	Swimmers	200	27	32	.00	4.6
45	Track	100	25	30	—.25	3.7
10	Bk 1937	73	27	31	—.14	3.7
29	Crew 1948	47	29	32	—2.00	4.5
12	Crew 1936	98	30	34	—.54	3.5
188		518			—.35	4.1

tests taken at random suggests that the deviations from it are largely due to emotion.

A male college athlete sitting at rest should have a flow index less than 25 and it should not rise above 30 after standard stimulus. Yet more than half of them do, probably because of emotion. In fact, the flow index is often observed to be twice the expected, and it may fall, not rise, after standard stimulus. It is perfectly evident that there is no physiologic need for such erratic changes. They may be due to vasomotor reactions which themselves are largely due to emotion. Some champions have shown no such vagaries in tests over their whole college careers.

The effect of emotion can be differentiated from the physical changes. Thus a fall in the index after standard stimulus, when a rise of five units is expected, can mean only that the heart was working quite unnecessarily at rest before exercise, and calmed down, as the result of the stimulus, more than five units. A high F to begin with reinforces the argument. On the other hand, a flow index F = 20 or less which increases more than five after standard stimulus shows that the subject was truly at rest before exercise, but the stimulus produced an excessive response. This may be merely

emotional. To prove it requires success in dispelling the emotion, which is not always easy. The relation of emotion to impairment is still further involved.

Discussion

A very real oxygen debt is incurred in athletic events. Shortly after a workout, however, it is not unusual to observe a flow index of 20 or less, long before the oxygen debt has been cleaned up. Apparently, the circulation is slow in working it off. So the immediate response to exercise is largely anticipatory in nature. Indeed, the mere thought of a contest is enough to set the heart racing. As soon as the exercise ceases, the threat to the athlete is removed and he relaxes to a comfortable flow index. Thus, performance tests are inextricably involved with emotion.

Standard stimulus has been very carefully designed to avoid emotional stimulus. If the same number of ascents is performed in less than 90 seconds, the responses are greater. Any novelty, such as attaching electrocardiographic leads, produces an emotional reaction even in those complacent to the standard technique. Ordinarily, the flow index (F') decreases from test to test, as familiarity lessens emotion. The fact that much can be done to induce an athlete to relax is perhaps the best evidence that the vagaries in circulatory response are largely due to emotion.

To illustrate by examples requires too much detail; but an experience of 20 years testing 426 athletes repeatedly, many through their whole college careers and some even longer, has convinced the author that the flow index is the best practical way to discount the emotional factor in flarimeter tests.[3]

REFERENCES

1. Dawson, P. M., *Physiology of Physical Education*, Baltimore: Williams and Wilkins Co., 1935.
2. Wells, P. V., *Flarimeter Manual*, printed by Princeton University Press, 1953.
3. ———, An improved flarimeter, *American Heart J.*, **47**: 102-104, 1954.
4. ———, Flarimeter tests on athletes, *American Heart J.*, **48**: 914-928, 1954.
5. Wiggers, C. J., Physiology from 1900 to 1920, *Annual Rev. Physiol.*, **13**: 8, 1951.

[3]I wish to thank the Princeton University Department of Health and Physical Education for their courtesy and help at every stage, the men who graciously volunteered for the tests, and the coaches, especially Howie Stepp and Matty Geis, for their co-operation.

(Submitted 9/28/54)

Notes and Comments

NOTES

Professional Courses in the Prevention and Treatment of Athletic Injuries at American Universities and Colleges

LANE B. BLANK
Franklin Junior High School
Long Beach, California

THE PURPOSE OF this study was to examine and analyze the course offerings in the prevention and treatment of athletic injuries for the preparation of male professional personnel in physical education at American universities and colleges.

The first portion of the study consisted of the examination of the catalogs of all American institutions of higher learning which offer professional education in the fields of health education, physical education, and recreation. Of the 419 schools investigated 222 were discovered to offer the type of course in question, and 13 of the latter were found to have two such courses.

The principal technique of investigation consisted of mailing questionnaires concerning specific courses and their characteristics to each of the aforementioned 222 institutions. One hundred twenty responses (approximately 54 per cent) were received, thus constituting a fairly adequate degree of response (1).

The findings of the present study are summarized as follows:[1]

1. More than one-half of the courses in athletic injuries are classified at the level of the third-year college student.

2. The following titles account for more than 60 per cent of the courses investigated: "prevention and treatment of athletic injuries," "athletic training and conditioning," and "first aid and care of athletic injuries."

3. The most common prerequisites for courses in athletic injuries appear to be (in order): anatomy, physiology, first aid, kinesiology, and physiology of exercise.

4. A course in athletic injuries is required for the attainment of the Bachelor's degree or teaching credential in physical education at the majority of institutions.

5. Nearly 70 per cent of the courses were reported to be benefited by the appearance of guest instructional personnel, particularly team physicians and athletic trainers.

6. Study of the prevention, care, and treatment of athletic injuries ranks the highest in regard to the amount of class time consumed, while the least time is devoted to the responsibilities and limitations of the trainer and coach and to the study of training facilities and equipment.

7. The most widely used teaching techniques employed appear to be the lecture, demonstration, laboratory, and examination activities, while the least widely used experiences include group meetings and reports, field trips, visits, and observations.

8. The most popular textbook in use at present is that of Cramer (2).

9. The more popular items of subject matter suggested by respondents for future inclusion in courses in athletic injuries are nutrition, the specifics of treatment, physical therapy, and review of the basic sciences.

REFERENCES

1. American Association for Health, Physical Education, and Recreation. *Research Methods Applied to Health, Physical Education, and Recreation.* Wash., D. C.: AAHPER, 1952. 535 pp.
2. Cramer, Frank L., L. L. Boughton, and Charles Cramer. *A Training Room Manual.* Gardner, Kans.: Cramer Chemical Co., 1945. 144 pp.
3. Herrington, George Squires. An analysis of courses in educational sociology with proposed changes. *J. Educational Sociology* **22**: 259-75 (Dec. 1948).
4. Stafford, Frank S., and Stella T. Sebern. *Institutions Offering Professional Education in Physical Education.* Wash., D. C.: U. S. Govt. Printing Office, 1949, pp. 1-4.

[1]Copies of the tabular matter summarized herein will be supplied by the author upon request.

Predictive Ability of a Selected Basketball Test

CLIFFORD A. BOYD, JAMES R. McCACHREN, AND I. F. WAGLOW
University of Florida
Gainesville, Florida

ONE OF THE PROBLEMS confronting a coach is that of economy of time and effort in making an efficient selection of his squad members. With this in mind, the investigators decided to determine the predictive ability of a test battery in selecting members of a basketball team. A review of the literature revealed the Knox Basketball Test which reported a reliability coefficient of .88 as a good battery to use. This battery consisted of a speed dribble, wall bounce, dribble and shoot, and a "penny-cup" test.

The test was administered to 42 candidates for the University of Florida Junior Varsity squad. At the end of three weeks of practice, the coach was asked to rank the players according to playing ability. The coach retained the top 18 on his squad and the remaining 24 players were dropped. The coach's ranking was placed in a sealed envelope to be opened at the end of the season. At the end of the season, the coach also ranked his 18 squad members as to ability without consulting any of the written records.

In order to arrive at an equitable system of ranking players regarding playing time, the total number of minutes played was divided by the number of games played, thus yielding an average playing time per game. The players were also ranked by comparing total points scored with average minutes played per game.

A bi-serial correlation of .96 was revealed when the test results were correlated with the categories of team membership and non-team membership.

The five variables were correlated using the rank difference method of correlation. Results are given in Table 1.

TABLE 1

Intercorrelations of Knox Basketball Test, Coach's Rank (End of 3 Weeks and End of Season), and Game Factors (Minutes Played and Points Scored)

Variables	1	2	3	4	5
1. Test	X	.17	−.10	.00	.11
2. Coach (End of Three Weeks)		X	.27	.23	.52
3. Coach (End of Season)			X	.85	.79
4. Average number of minutes played per game				X	.11
5. Total points scored compared with average number of minutes played per game					X

There were only two significant rank difference correlations. First, a rho of .85 was obtained between the coach's ranking of the 18 squad members at the end of the season and the ranking of the 18 players according to average minutes played per game. Secondly, a rho of .79 was obtained between the coach's ranking of the 18 players at the end of the season and the ranking of these players as to total points scored per minute of playing time. This would indicate that the coach made a valid rating of his players if time played and the number of points made per minute of playing can be considered as valid criteria of playing ability.

The Knox Basketball Test appears to have merit as a predictor of squad membership in basketball; that is, it can readily distinguish between varsity and non-varsity caliber players. On the other hand, the value of the Knox Test in distinguishing between levels of ability among squad members does not have any statistical significance.

REFERENCES

1. CLARKE, H. HARRISON. *Application of Measurement to Health and Physical Education.* New York: Prentice-Hall, Inc., 1950. pp. 305-306.
2. EBEL, E. R., and FORREST C. ALLEN, Evaluating team and individual performance in basketball. *Research Quarterly,* 5: 538-555, 1941.
3. EDGREN, H. D., An experiment in the testing of ability and progress in basketball. *Research Quarterly,* 3: 159-171, 1932.
4. FRIERMOOD, H. T., Basketball Progress Tests Adaptable to Class Use. *J. Health and Physical Education,* 5: 45-47, 1934.

5. GARRETT, HENRY E. *Statistics in Psychology and Education.* New York: Longmans, Green & Co., 1947. pp. 343-353.
6. KNOX, ROBERT D., Basketball ability test. *Scholastic Coach,* 17: 45, 1947.
7. LEHSTEN, NELSON, A measure of basketball skills in high school boys. *The Physical Educator,* 5: 103-109, 1948.
8. VOLTMER, E. F., and TED WATTS, A rating scale for player performance in basketball. *J. Health and Physical Education,* 2: 94-95, 1940.

COMMENTS

Comments on the article by Alfred W. Hubbard and Charles N. Seng, "Visual Movements of Batters," *Research Quarterly,* 25: 42-57 (March 1954).

ALTHOUGH ONE continues to hear echoes of the attractive Baconian notion that sciences proceed best by avoiding all hypotheses and beginning with the facts themselves, there has been a growing recognition of the importance of hypotheses in any scientific enterprise. Hypotheses constitute the first step in any scientific undertaking. The particular ones proposed for a problem set the stage for observation. Without them one would not know what to observe. And the observations serve to confirm or reject the hypotheses involved.

Occasionally, however, investigators tend to forget the role of hypotheses in the scientific enterprise, and they behave as though their hypotheses were established facts.

In their report on visual movements of batters, Hubbard and Seng state that:

". . . batting is not primarily a reaction-time problem. The stimulus object (ball) is continuously visible during its flight, not suddenly presented. Consequently the problem is one of tracking a moving object, predicting its course and, at some point in its flight, deciding to swing or not. This prediction subserves the motor response. Although generating momentum in the bat, and the preceding central processes, involve some duration, any sensory component in reaction time is essentially eliminated. In other words, estimates based on reaction time or when the batter must make his decision and start the process which will result in the swing place the point too far back in terms of ball flight." (p. 42)

Their analysis presents two interesting hypotheses relative to the general batting situation. One hypothesis is that the batter can track a pitched ball, and the other is that the batting response is unlike a reaction-time response because the sensory component is essentially eliminated. As is the case with all hypotheses, one can and should ask the question, Does the evidence confirm or refute?

The main evidence investigators offer for tracking a pitched ball is presented in Table 5, entitled, "Incidence of Head and Eye Movements by Frames from Complete Sequences in Reaction to Flight and Speed of Ball." (p. 51) Inspection of these results reveals that no apparent tracking movements were found during practically the first half of ball flight. In short, there was no evidence of tracking during what may be considered the critical period of ball flight. The investigators, of course, are willing to concede that their measurement technique probably was not adequate for recording possible head or eye movements during the early part of flight, but this limitation cannot be said to constitute a guarantee that tracking did occur.

If one considers the interval of ball flight up to the point where the central processes took over and led to the generation of momentum in the bat, as the investigators explain events, the situation is not much better. It is stated, for example, that the momentum for the swing was probably generated and applied to the bat during the fourth or fifth frame before the hit. If one then used the extremely conservative estimate of two frames (approximately 0.08 sec.) for the preceding central processes, making an over-all estimate of six or seven frames, Table 5 is indeed revealing. Practically all instances of apparent head and eye movements occur after commencement of the central processes. Since the central processes presumably cannot do much in the way of predicting the ball's course, deciding to swing or not to swing, and generating momentum in the bat until information has been received, any evidence of tracking is essentially eliminated.

The principal thing the investigators have to show for their efforts is that there appeared to be some apparent head and eye movements after it was too late to track the ball. Precisely how any conclusions relative to tracking can be reached under these conditions is difficult to understand. The fact that the investigators were able to even specify the type of eye movement batters are supposed to use in tracking can only indicate an act of supreme faith in the ability of batters to track a ball. In other words, the hypothesis of tracking was assumed to be a fact at the outset, and interpretations of the data followed the implications of this assumption. To state that such procedures seldom prove fruitful is a mere truism.

Although the limitation of a note does not permit a detailed evaluation of the hypothesis that batting is not primarily a reaction-time problem, it is possible to comment briefly on one of the consequences involved. The investigators, for example, object to the reaction-time hypothesis on grounds that "estimates based on reaction time of when the batter must make his decision and start the process which will result in the swing place the point too far back in terms of ball flight." On the basis of this assumption, one should expect estimates of the interval between commencement of the central processes and the hit to be shorter than reaction time estimates of batting. That such is probably not the case can be simply demonstrated.

Using the conservative estimate of six or seven frames before the hit as the point where the central processes must begin to act on information previously received, and using a time interval of 0.04 sec. between frames, it would seem that the investigators' batters must have started reacting to the ball's flight no later than 0.24-0.28 sec. before the hit. In contrast to this interval, studies on batting reaction time found that starting reaction times in simple and choice situations required 0.21-0.29 sec. The surprisingly close agreement between these two estimates can hardly be said to support the hypothesis involved.

It is to be emphasized that the estimate of 0.08 sec. for the central processes is undoubtedly far too conservative. Even in a relatively simple response situation, the central processes appear to require the greatest proportion of time. For example, in his discussion of a tracking situation involving a reaction time of approximately 0.30 sec., Craik (1) notes that:

> Physiological studies of the response time of the eye and the time taken for impulses to be transmitted up the optic nerve show that these can account for about 0.01 sec. only, and electrical stimulation of motor nerves and records of the start of the limb-movement which results show that this accounts only for another 0.01 sec. approximately. There is left about 0.28 sec. to be accounted for in the cerebral processes.

REFERENCES

1. Craik, K. J. W. Theory of the human operator in control systems. II. Man as an element in control systems. *British J. Psychology*, 38: 142-148, 1948.

—*Arthur T. Slater-Hammel, Indiana University.*

Rebuttal to above comments on "Visual Movements of Batters"

SLATER-HAMMEL CONJURES up two hypotheses from a study which his introductory statements lead one to believe had none. The actual hypotheses established for testing may be found in the "Purpose of the Study," the preceding paragraph and various parts of the text (p. 43 ff.). But since he chose to set up as "hypotheses" an assumption and an opinion, neither of which was established for testing, we shall discuss them.

Slater-Hammel's first "hypothesis" concerns our assumption that "a batter can track a pitched ball," or that visual perception depends on tracking the ball. Distinct vision requires a stationary image on the retina. So, in batting, eye movement consistent with ball movement (up to the limits of eye movement) would seem necessary to subserve perception of the speed and direction of pitched balls. One could tell pretty well whether the eyes were focused on the ball in frames in which both were visible, but evidence of tracking before that necessarily depended on eye movement consistent with

ball movement. Furthermore, evidence of eye movement inside 30 ft. presupposed previous tracking. In interpreting our Table 5, the important point was the *absence* of evidence of tracking in the last frame or two, rather than lack of evidence in early frames (p. 51).

The absence of evidence of tracking in early frames is inherent in the situation and integrated with the perceptual part of the batter's problem—to differentiate speed and direction of pitched balls. In solving this problem preliminary movements by the pitcher may be misleading, but after delivery the batter's cues come from increase in apparent ball size and ball movement. Differences in the rate of increase of apparent ball size probably provide cues to ball speed. But we were more interested in apparent ball movement, which can be reduced to horizontal and vertical movement—the latter depending on angle of projection and ball speed, which complicates computation. However, assuming ball flight in a vertical plane, the *horizontal* angular movement would be the angle whose sine is proportionate movement divided by distance from the batter. That is, for a ball that would pass in a vertical plane 6 in. from the batter's eyes, the horizontal angular movement from release to 30 ft. from the batter would be 3/360, or 29′ (minutes), and for a ball passing 36 in. from him, 18/360, or 2°51′. To put it briefly: half of the apparent horizontal movement from release to 30 ft. occurs from 40-30 ft.; as much apparent movement occurs between 30 and 20 ft. as between release and 30 ft., and about three times as much occurs between 20 and 10 ft. as between release and 30 ft. or between 30 and 20 ft.

Note the small apparent angular movement from release to 30 ft. or even to 20 ft. For the batter, differences in the rate of angular movement and increase in ball size during the first half of ball flight provide a poor basis for differentiating ball direction and speed, but inside 30 ft. the bases become exponentially better. For the experimenters, such slight corresponding eye movement during the first 30 or 40 ft. of ball travel spread over 5 to 12 frames was not measurable. Microscopic analysis of enlarged prints might have shown evidence of movement. But since evidence of tracking inside 30 or 20 ft. presupposed previous tracking and since one of the primary problems was to determine whether the ball was tracked *to contact*, the additional labor and expense seemed unwarranted. Failure to find evidence of tracking preceding that in Table 5 was, as we explained, an artifact of the method and situation. It should not be interpreted as evidence that tracking was *not* occurring. If such interpretation was unfruitful, so be it; it was consistent, reasonable, and necessary.

We disagree emphatically with the conclusion that, "In short, there was no evidence of tracking during what may be considered the critical period of ball flight." Slater-Hammel arrives at this by discarding all evidence of tracking after some hypothetical point at which the central process must "start." The central processes would seem to be going before, during and after the swing, rather than being forced to start at some prearranged signal. A temporal limit must exist before which the batter must make his choice to start the swing—or accept the consequences. But after starting the bat he may still elect not to complete the swing. This action occurs at about the time the ball reaches the plate, suggesting that he can use cues collected as or after the swing starts. If he can use these late cues to check his swing, he can presumably also use these late cues to attempt to direct the bat to contact; but he must collect them before the angular movement of the ball exceeds the limits of eye movements and before much momentum is developed in the bat. Analysis of the rate of change of apparent ball size and movement suggests strongly that the first 30 or 40 ft. provide cues much less easily differentiated than the last 30 or 20 ft. (except for the last 5 or 10 ft.). Table 5 shows evidence of tracking in this latter period so we see no need to write off the evidence on the basis of invalid assumptions.

Now concerning Slater-Hammel's second "hypothesis": "that the batting response is unlike the reaction-time responses because the sensory component is essentially eliminated." Batting and batting reaction-time are not analogous because in batting one must

respond differentially to a serial stimulus situation as it develops—with certain limiting factors. In batting reaction-time the subject responded to a suddenly presented prearranged signal by starting the bat along a track. Sports situations which force the athlete to respond to a suddenly presented stimulus for which he is set are analogous to the reaction-time situation. But situations in which he can see the elements develop and to which he must respond differentially as best he can seem to be perceptual (and motor) problems. A more serious objection to the batting reaction-time approach arose from an analysis of that study.

Let us assume that batting can be studied as a reaction-time problem since, "In batting, a player must react to a ball in flight, and it would seem that a batting reaction-time factor must enter."[1] This logic seems highly questionable since it implies that every time a person reacts a reaction-time *must* enter. However, we will accept a mean of .21 sec. for "starting reaction-time" and of .27 sec. for "movement reaction-time." Then, ". . . it requires some time to bring the bat forward after starting the movement, and this would have to be added to the reaction time" (p. 356). The interval for moving the bat forward seems to be (for 86 of 96 cases) about .12 to .20 sec., or on the average about .16 sec. By addition we get, using mean figures, either .37 or .43 sec., or using .12 and .20 sec., between .33 and .49. According to the rules, these estimates of the interval to "react and move" should have been used in going from reaction-time to the batting situation.

To find the point at which the batter would have to get his signal to start to react, we multiply ball flight (60 ft. roughly) by the fraction, interval to "react and move" divided by interval of ball flight at the speeds of pitched balls. A ball travels 60 ft. at 50 mph. in .82 sec., at 70 mph. in .58 sec., at 95 mph. in .43 sec., and at 100 mph. in .41 sec. Using mean figures for "starting reaction-time" plus "movement-time," we get for 50 mph. 32 ft., for 70 mph. 45 ft., for 95 mph. 52 ft. and for 100 mph. 54 ft. Using mean figures for "movement reaction-time" plus "movement-time," we get corresponding estimates of 32, 45, 60, and 63 ft. Using .20 sec. "movement-time," the corresponding distances for "starting reaction-time" would be 30, 42, 57, and 60 ft., and for "movement reaction-time" 34, 49, 65, and 69 ft. Note that for fast balls or for something like a "movement reaction-time" the batter may have to get his signal *as or before the pitcher releases the ball.*

Following the specified rules for determining when the signal must occur in assuming batting is a reaction-time situation, we find: first, that only for very slow balls could the signal occur inside the last half of ball flight; and second, that the batter must be prepared to react and swing on the basis of some signal which *may* occur before the pitcher releases the ball. Since he cannot know until release by the pitcher what the situation will be and since the first half of ball flight provides very poor cues, how can he differentiate between even a fast and slow ball, let alone estimate its direction? We found (in Table 2) that the batters did respond differently to balls at different speeds. They responded differentially to balls arriving in different positions—by hitting them. Consequently, our opinion that batting is a perceptual (and motor) problem rather than something analogous to reaction-time seemed valid.

Only by disregarding movement-time can one get the estimates of "22 ft. to 30 ft. from home base" for a "starting reaction-time" to balls 70-95 mph., or "28 ft. to 38 ft." for a "movement reaction-time" (p. 355). The proponent of ". . . surprisingly close agreement . . ." chooses to perpetuate the error of disregarding movement-time. For us, "batting reaction-time" remains a primarily academic problem with little or no demonstrable analogy to actual batting. In other words, there must be a better approach to the problem of batting than one which disregards the facts of batting and appears to provide no feasible basis for differentiating ball speed and direction.—*Alfred W. Hubbard, University of Illinois.*

---

[1]A. T. Slater-Hammel and R. L. Stumpner, Batting reaction time. *Research Quarterly*, **21**: 353 (Dec. 1950).

Research Abstracts

Prepared by the Research Abstracts Committee of the National Council of the Research Section, PAUL A. HUNSICKER, Chairman

Health

45. National Dairy Council. Adolescent nutrition. *Dairy Council Digests,* **26:** 4 (Mar. 1955).

The recommended calorie allowances for boys and girls are differentiated at age ten and subsequently through adolescent and mature years. At ten, at least in the United States, boys are more active and the same activity seems to cost a boy more energy than it costs a girl of the same size.

Special attention needs to be given to providing the adolescent with enough calories, protein, calcium, vitamin D, and iodine. Generally, if foods are included which supply these nutrients, other nutrient needs will also be met.

The energy and nutrient needs of adolescents are greater than for other periods of life with the exception of pregnancy and lactation in women.—*Paul Hunsicker.*

46. Eichenlaub, John E. Cataract. *Today's Health,* **33:** 32-33 (Mar. 1955).

Cataract is but a fogging of the lens in the eye. Difficulty with vision, especially for close work, and a faint grayness of the pupils are usually the first symptoms noticed. If the eyes are left alone, this grayness of the pupil will ultimately turn a dead white, the lusterless white of a ripe cataract. Except for this whiteness, there is no sign of trouble—no inflammation, no growth of tissues, no irritation of the lids. Many cataracts start in the middle part of the lens because this part receives the least nourishment. Some babies are born with cataracts, the lens never having developed properly. Diabetes and glaucoma cause changes in the chemistry or circulation of the eye fluid and speed the formation of cataracts. There is only one cure—the clouded lens must be removed. Such operations under local anesthesia have proven amazingly successful under competent eye surgeons.—*J. Grove Wolf.*

47. Eichenlaub, John E. Sciatica. *Today's Health,* **33:** 23: 40-43 (Apr. 1955).

Sciatica doesn't kill, doesn't take away your ability to move, doesn't make you weak, but it does disable you. Most medical authorities agree that the cause of sciatica is usually rupture of an intervertebral disc. There is a difference of opinion even among top authorities on the proper way to handle a ruptured disc. There will be some disability after an operation, so one should not get out the knife too soon. Probably the best single test to show if a ruptured intervertebral disc needs an operation is the straight-leg-raising test. One noted authority will not operate on anyone who can lie on his back and raise the affected leg more than halfway off the table without bending the knee. One may help prevent sciatica by (1) avoiding lingering mild infections; (2) lifting properly without bending the back—lift with legs, not spine; (3) keeping posture normal; and (4) not greeting every backache with a softer and softer mattress.—*J. Grove Wolf.*

48. Eppright, E. S., and Charlotte Roderuck. Diet and nutritional status of Iowa school children. *Amer. J. of Public Health.* **45:** 4: 464-471 (Apr. 1955).

The food intake and nutritional status of children were studied in Iowa. There were 61 schools considered with a sampling of 1,200 students. The core of diets of Iowa children, like their elders, consisted of liberal amounts of meats, potatoes, sweet desserts, table fat, bread, and cereal. The milk consumption was two to three cups daily. Diets were slightly low in vitamins C and A. The most significant problems were dental caries and weight control for the teen-age girls. Iowa children need only slight improvements in their diets to obtain good nutrition.—*Joseph Lantagne.*

49. JOHNS, EDWARD B. Field experiences in health education for prospective teachers. *J. of School Health*, **25**: 5: 123-131 (May 1955).

The preparation of teachers in health education is one of the major objectives facing school health education today. In the preparation of health educators are these persons obtaining the kinds of field experiences necessary to meet the challenges they are about to face? Dr. Johns summarizes by stating "First, the need for purposeful experiences in the preparation for teachers in health education; secondly, the meaning of laboratory experiences, including field experiences, and desirable outcomes; thirdly, practices and current problems relating to laboratory experiences in the preparation of school health educators.—*Joseph Lantagne.*

50. SALK, JONAS E. Present status of the problem of vaccination poliomyelitis. *Amer. J. of Public Health*, **45**: 3: 285-297 (Mar. 1955).

There are many who hold that an effective means of preventing paralytic poliomyelitis will come from a live-virus vaccine. This live-virus vaccine should be more desirable even if a non-living vaccine is found to be effective. Life-long immunity seems to be associated with recovery from the naturally acquired infection. Most scientists agree that it isn't possible to reproduce the degree nor the quality of immunity from non-living virus that is attainable from living virus. Results from experiments indicate that, even with a non-living vaccine, by proper spacing and dosage, it will produce a booster effect. From these data either a living or non-living vaccine should produce desirable results.—*Joseph Lantagne.*

51. TOBEY, JAMES A. Backache. *Today's Health*, **33**: 42-43 (Feb. 1955).

Over 57 possible causes of backache have been listed by various authorities including poor posture, sudden strains, arthritic conditions, and many others. Pains in the back are almost invariable accompaniments of many of the acute communicable diseases such as influenza, pneumonia, and other acute respiratory disorders. The numerous physical causes of backache are estimated to account for only a little more than half the poor aching backs with which the American people are afflicted. The remainder are not due to physical causes, but arise from deep-seated nervous and emotional disorders. It is much better to prevent the affliction of aching backs than to effectively treat this condition. This can generally be accomplished by the maintenance of general good health; by proper posture; by avoiding overweight; by preventing strains and accidents; by dodging contagions; and by cultivation of good mental hygiene.—*J. Grove Wolf.*

52. WITTY, PAUL, Comics, television and our children. *Today's Health*, **33**: 18-21 (Feb. 1955).

Over 90 per cent of boys and girls between eight and 13 years of age read comic books regularly. Fourth-, fifth-, and sixth-graders appear to be the most avid readers. The attraction of comic magazines continues throughout the middle grades, and many high school students read comics.

Children spend nearly 20 hours a week watching television during the year, with the high school student spending somewhat less time, averaging between 14 to 17 hours a week. No significant differences were found between the IQ's of pupils and the hours devoted to TV. Excessive viewing of television, however, seemed to be associated with somewhat lower academic attainment. The teachers concluded that the desirability or undesirability of the amount of televiewing could be determined only by a complete study of each child. Parents and children should together plan balanced programs of recreation not only with time for children to see TV and read comic books, but also with time for them to engage in other activities of greater value.—*J. Grove Wolf.*

Physiology

53. BLANK, LANE B. Effects of warm-up on speed. *Athletic J.* **35**: 10: 45-46 (Feb. 1955).

A group of 16 experienced track and field athletes were divided at random into two groups and timed daily in the 120-yard dash for 22 consecutive days. One group par-

ticipated in time trials following an optimum warm-up pattern, while the other group underwent similar recorded runs preceded by a minimum pattern of activitiy. The warm-up routines were interchanged on alternate days. In a second experiment, 48 individuals with no prior track and field experience volunteered as subjects, but only 38 completed the research period. Again the alternating pattern was employed, but the distance run was reduced to 100 yards. The difference between the mean optimum warm-up and mean minimum warm-up for each of the two groups was found to be significant at better than the one-tenth of one per cent level of confidence. It was found that warm-up is of benefit to performance in sprinting and for the groups investigated there was no apparent difference between the effect of the warm-up routines upon experienced and non-experienced individuals.—*J. Grove Wolf.*

54. Dill, D. B., J. C. Seed and F. N. Marzulli. Energy expenditure in bicycle riding. *J. Applied Physiol.* **7**: 320. (1954).

Energy expenditure in riding a bicycle with balloon tires (2½-in. diam.) and narrow tires (1½-in. diam.) was made. The net oxygen consumption was 0.19 L/min. greater with balloon tires. This difference is explained by a greater friction caused by the balloon tire. At a speed of 15 mph the propelling force for the balloon tire is 6.5 gm/Kg. gross load, whereas for the narrow tire it is 2.5 gm/Kg. Bicycles were ridden on an electrically driven treadmill.—*Leslie Leggett* and *Peter V. Karpovich.*

55. Gersten, J. W. Changes in hydration of muscle and tendon following the application of ultrasonic energy. *Arch. Phys. Med.* **36**: 140-148 (Mar. 1955).

Exposure to ultrasonic vibration increased the relative water content of 24 normal frog gastrocnemii, while decreasing that of 30 normal Achilles tendons. The changes were reversible. Since anoxia, temperature elevation and 2, 4-dinitrophenol (DNP) produced little change in the dehydration effects elicited by ultrasound, the author suggested that physical factors were of primary importance in controlling the water content. This hypothesis was given support by the changes in the hydration occurring after treatment with glycerol or 0.5M KCl, in which case increased hydration was observed in muscle and also in tendon.—*Clifford E. Keeney* and *Peter V. Karpovich.*

56. Hellebrandt, Frances A. Physiology of motor learning as applied to the treatment of the cerebral palsied. *Qaurt. Rev. Pediat.* **7**: 5-14. (1952).

The degree of neuromuscular training affects all phases of the total rehabilitation program. The basic motor patterns which are not under volitional control, as in the cerebral palsied individual, can be controlled by training.

Attainment of complex motor acts by utilizing individual muscles or isolated muscle groups may be unphysiologic. Motor skills appear to develop out of total patterns of response. As overloading is important for the development of power so excessive stimulation may be just as important in acquiring motor skills. Patient should be positioned to permit functioning of normal developmental patterns, rather than placed in a position providing for only isolated muscular movements.—*Elena M. Sliepcevich.*

57. Henry, F. M. Aerobic oxygen consumption and alactic debt in muscular work. *J. Applied Physiol.*, **3**: 427-438 (Jan. 1951).

Oxygen intake during exercise and recovery was measured continuously while 12 subjects worked on a bicycle ergometer. The load ranged from 15 to 910 kg. m/min at a speed of 62 r.p.m. The steady state oxygen intake a_o was linearly proportional to metabolic work load, and there were no individual differences in intake rate. The intake rose to the steady-state value in accord with an exponential law, the velocity constant k of this rise being correlated $r = 0.79$ with the alactic recovery constant k', and exhibiting pronounced individual differences.

In another experiment, 35 subjects worked at a constant load of 614 kg. m/min. It was found that steady-state intake a_o and the recovery exponent k' were uncorrelated. Athletic training did not influence a_o but caused a significant increase in k'. All of the results were explained theoretically by an exponential mathematical model which states

that oxidizable substrate is produced linearly in proportion to the work load, and is continuously oxidized, so that the steady state represents a balance between substrate production and oxidation. The initial lag of intake at the beginning of exercise is not due to slow adjustment of the cardiovascular system, but is a necessary consequence of the fact that substrate cannot be oxidized until it has first been produced by the muscular work.—*F. Henry.*

58. Kaplan, Harry A., and Jefferson Browder. Observations on the clinical and brain wave patterns in professional boxers. *J. of Am. Med. Assn.*, **156:** 1138-1144 (Nov. 20, 1954).

The authors have conducted a study, over a period of four years, of 1,043 professional fighters with particular reference to responses resulting from contact trauma to the head inflicted in the boxing arena. One or more electroencephalograms was made for each subject, the total number of such examinations for the entire study being 1,400. The study was carried out in three parts: clinical observations at ringside, cinematography, and electroencephalography.

Observations at the ringside and more detailed examination in the dressing room after the fight failed to reveal any abnormal neurological features, even in those contenders who lost their bouts by a knockout. A critical review of slow-motion photography of segments of fights confirmed one major impression gained at the ringside, namely, that most blows to the head were short of their mark or deflected by the opponent. Electroencephalographic findings were classified according to 12 types of patterns. The greatest percentage of fighters was found in the slightly slow anterior type of pattern, consisting mainly of moderate voltage, 5 to 7 per sec. frequency, but with a fair amount of low voltage beta and alpha activity intermixed and superimposed in the frontal region and constant well-formed 10 per sec. alpha posteriorly. The physical features and performance data of each fighter were then tabulated for correlation with the electroencephalogram. Specifically, the age of the contestant, the number of fights, the rating of the fighter in relation to the title holder, the type of fighter, the weight division, the number of wins or losses, and the number of losses by technical or 10 sec. knockouts were the items used in the correlations. Results failed to reveal any significant statistical findings, except in the rating class, in which statistical results indicated that those lower in ring rating had the greater percentage of disorganized electroencephalograms. The top rated fighters were found most prominently in the low voltage and fast types.

From one to six repeat electroencephalographic examinations were carried out on 197 fighters. While changes occasionally were noted in the records of some subjects, there was no uniform pattern of change with continuity, sytle, or frequency of fighting. Many fighters after one, two, and three years of active boxing had no change whatsoever in their electroencephalographic tracing. Even the records taken on 40 fighters within 10 minutes of the cessation of a bout showed no change over their prefight record. The authors believe that the electroencephalographic pattern of professional fighters, as well as members of the general populace, may be more inherent in the person than secondary to occupational pursuit.—*Digest of Neurology and Psychiatry* (Dec. 1954).

59. Koepke, G. H., A. J. Murphy, J. W. Rae, Jr., and D. G. Dickinson. An electromyographic study of some muscles used in respiration, *Arch. Physical Med.*, **36:** 217-222 (Apr. 1955).

Respiratory movements of the chest were recorded by means of a pneumograph attached to a pressure transducer. The action potentials were picked up by needles and recorded by a single channel electromyograph. Forty-five men were used as subjects.

During a quiet inspiration only the first and the second intercostal muscles were active. During a deep inspiration intercostal muscles up to and including the seventh were used. It was difficult to differentiate between the internal and external intercostals. The abdominals were more active during expiration than inspiration. During a maximum expiration, obliques were used more than the rectus.—*Peter V. Karpovich.*

60. Zankel, H. T., R. A. Shipley and R. E. Clark. Effect of posture and local pressure on venous circulation time of the leg. *Arch. Physical Med.*, **36**: 226-231 (Apr. 1955).

Circulation time was determined by injecting radio-iodine (I^{131}) into the dorsal vein of the foot and detecting the appearance of I^{131} in the skin over the groin by many a scintillation counter.

The average time for 24 patients was 28 seconds, with a range of from 8 to 90 sec. Application of pressure to popliteal region reduced the circulation time to 14 sec. with a range of from 3 to 50 sec. Pressure of 10 to 20 mm applied to calf reduced the circulation time to 21 sec. Flexion of the knee and assumption of a sitting position produced no consistent effect. Time was increased on standing position. Elevation of legs decreased the circulatory time. In patients with arteriosclerosis, obliterance, elevation of legs caused an increase in circulation time.—*Peter V. Karpovich.*

Psychology

61. Bernberg, Raymond E. A measure of social conformity. *J. of Psych.*, **39**: 89-96 (Jan. 1955).

Two forms of an attitude scale to measure social conformity were constructed, using items classified in such categories as moral values, positive goals, reality testing, ability to give affection, tension level, and impulsivity. The direction of perception technique of altitude measurement was used. These scales were initially given to two groups of boys: 100 high school seniors of low socio-economic level and low IQ, and 125 boys from a broad range of socio-economic level, average IQ of 114 on Otis short form.

A scale (the HRI) was revised and to secure evidence of discrimination between groups was given to: (a) 160 male inmates of a California youth prison, (b) 101 Police Science majors. Significant mean differences between conforming and non-conforming groups were found by both parametric and non-parametric tests of significance. The scale was further checked to determine the influence of intelligence, socio-economic status, cultural background (sex, age within limits, and religion) on the responses. These factors were found insignificant in their influence on the responses.

The authors recommend this scale as a research tool, as a tool for screening where social conformity reflects desirable traits, and as a demonstration tool of socio-psychological concept and indirect attitude measurement method.—*Frances Z. Cumbee.*

62. Feintuch, Alfred. Improving the employability and attitudes of the "difficult-to-place" persons. *Psycho. Monographs*, No. 392. **69**: 7 (1955).

The effect of a program which integrated vocational counseling, casework, and a sheltered workshop on employability was studied. Fifty-two adults in Montreal who had not been able to find and keep employment participated in the integrated program. Data on such items as time employed, attitudes toward work and self, and characteristics of the hard-to-place was collected prior to participating in the program and again one year after the program.

A comparison of the data before and after the program indicated the following results: (a) a significant increase in both employability and stability of employment following the program; (b) 7 of the 16 attitudes checked showed significant correlation with employability and that each of the 7 attitudes showed significant improvement during the program; and (c) 7 of the characteristics (age, dependents, able to get along well with people, disabilities, length of former employment, length of time in the country, and time getting financial assistance) differentiated significantly between high and low employability after the program. None of the attitudes seem to so differentiate.—*Frances Z. Cumbee.*

63. Henkin, Robert I. A factorial study of the components of music. *J. of Psych.* **39**: 161-181 (Jan. 1955).

The principal axis method of factoring was used on the response of students to records in an effort to isolate musical factors in esthetic appreciation. Ten records, chosen

from three periods (classic or pre-classic, romantic, and modern) to meet specifications under such categories as rhythm, melody, harmony, and color were used. The experiment was repeated twice with the same group of music appreciation students at UCLA; 243 taking the test the first time, and 193 taking the test the second time.

Six factors were obtained from the analysis, two of which the author felt justified in naming: (1) a *melodic factor*, and (2) a *rhythmic factor*. Some evidence was presented that the third factor might be an *Orchestral Color Factor*. The fourth factor was left unidentified and the author felt that the fifth and sixth factors were residual factors that could be used for information purposes in further studies. A study of factors not clearly defined was recommended in the hope that a clear definition of musical factors might be ascertained.—*Frances Z. Cumbee.*

64. Horst, Paul. A technique for the development of a differential prediction battery. *Psycho. Monographs.* No. 380. **68**: 9 (1954).

In a brief discussion of the differential prediction problem the need for a single classification battery of tests, differentially weighted, that would allow the prediction of success in each of a wide variety of activities was pointed out. The index of differential prediction efficiency for the battery was defined as a function of the variances of the predicted difference scores when all possible pairs of criterion variables were used. The larger the variance, the greater the differential prediction efficiency. Using the predicted criterion scores, it was shown further that mathematically an equivalent definition was obtainable by maximizing the difference between the average variance and the average co-variance.

An iterative procedure was used to develop an alternative technique that permitted progressively selecting the predictor which combined with the previously selected set would yield the largest index of differential prediction. This alternative technique was developed because the labor involved in exactly selecting a subset of predictors that would maximize the index would be prohibitive.

A numerical example was used to demonstrate the computational procedures. Still further, equations were developed for determining the matrix of regression vectors and multiple correlation coefficients without the necessity of calculating the inverse of the predictor intercorrelation matrix.—*Frances Z. Cumbee.*

65. Horst, Paul. A technique for the development of a multiple absolute prediction battery. *Psycho. Monographs.* No. 390. **69**:: 5 (1955).

A contrast between the multiple absolute prediction problem and the differential prediction problem was presented. It was pointed out that the multiple absolute prediction problem is one of selecting a subset of texts which would have the highest prediction efficiency for all criterion variables, regardless of how it differentiated among them. The sum of variances of the predicted criteria with no attention to the co-variance was used as the definition of the index of prediction efficiency.

Mathematical derivation of formulas for obtaining the matrix of beta weights and multiple correlations was presented. A numerical example is given which shows the computational procedure in use in selecting the absolute predictors, solving regression vectors and multiple correlations.—*Frances Z. Cumbee.*

STUDIES FOR CPEA REVIEW

Members of the profession who have recently completed their doctorates and would like to have their studies considered for possible review before the College Physical Education Association next year should contact *Charles Bucher, School of Education, New York University*, chairman of the Problems Committee for CPEA. Studies should concern problems of interest to people in college physical education.

Guide to Authors

IN LINE WITH the over-all goal of making Association publications yield the greatest value to the individual and the profession, the following is a guide for the preparation of manuscripts for the *Research Quarterly*, recognizing general techniques employed by research publications.

Article Manuscripts

Manuscripts should be sent to the Editor (AAHPER, 1201 Sixteenth Street, Northwest, Washington 6, D. C., who will see that each one is read by at least three members of the *Research Quarterly* Board of Associate Editors. On the basis of the three reviews, the Editor will advise the author as to the suitability of the paper or the desirability for revision. Papers are not judged by arbitrary standards but on their content of new research results in the field of physical education, health education, and recreation, presented with the greatest brevity compatible with scientific accuracy and clarity (see October 1951 *Quarterly*, pp. 392-4).

Since three members of the Board of Associate Editors review an article, it is requested that three clear copies of the manuscript be submitted in order to facilitate reviewing. A fourth copy of the article should be retained.

Typewritten manuscript should be double-spaced on white paper of ordinary weight and standard size (8½ x 11 inches). A brief abstract of the article, 100 words or less, should be typed double space on a separate sheet. See abstracts at head of *Quarterly* articles for style.

The sheets of manuscript should be kept flat and fastened with clips which can be removed easily. The pages of the typewritten copy should be numbered consecutively in the upper right-hand corner. Paragraphs should be numbered consecutively throughout the manuscript.

Notes and Comments

Notes on minor research and on apparatus, *objective* critical comments, and summaries of status suveys will be printed in the Notes and Comments section. Note that simple status surveys are no longer acceptable as regular *Quarterly* articles, by decision of the Research Council. Such studies will therefore be published in brief form (300-500 words) under Notes and Comments.

Headings

The article should be arranged so as to indicate relative values of heading and subheadings.

Usually four gradations are sufficient: (a) article title, (b) first subhead appearing in boldface aligned left on page (underscored in manuscript with wavy line) (c) second subhead (if necessary) appearing in small caps aligned left on page, (d) third subhead, to appear in italic (underscored in manuscript), not centered, but run in at the beginning of the paragraph or section.

All headings should be typed in lower case with initial capitals, except for (c) above, which should be typed in capital letters.

FOOTNOTES

Footnotes are not to be used for references or literature citations. They are rather used for the purpose of acknowledgment, special explanation, supplementary information, etc. (*See examples below.*)

Type footnotes (if any) on separate sheets, as many footnotes as convenient being written on a sheet. Footnotes should be numbered from 1 up for each article; a corresponding numeral appearing in the text. Asterisks should not be used.

Examples of Footnotes:

[1] This study was made under the direction of Dr. Arthur T. Slater-Hammel in the Research Laboratories, School of Health, Physical Education, and Recreation, Indiana University, Bloomington, Indiana.

[2] All measurements of the hand were recorded in centimeters and height was recorded in inches. The hand measurements were taken by Everett and reliability coefficients of above .90 were found for each measurement used in the study.

[3] For their wholehearted co-operation in facilitating collection of the data, special gratitude is extended to Superintendent Clarence Hines and the 1950-51 principals of the Adams, Condon, Edison, Francis Willare, Harris, Howard, Lincoln, River Road, and Whiteaker schools.

Documentation

CITATIONS OF LITERATURE

Citations of literature should be segregated alphabetically by author's last name at the end of each article, under the caption of "REFERENCES." *Do not treat them as footnotes.* (*See above.*)

The literature citations, listed alphabetically, should be numbered consecutively, their location in the text being indicated by corresponding numbers written in full size and enclosed in parentheses: for example, (1) (2, 3). If there are several references in the text to a citation, the specific pages may be indicated thus: (1, *p. 117*), (1, *pp. 162-3*).

A uniform style should be maintained in writing citations. Do not enclose titles of chapters and articles in quotation marks. Italicize (underscore in manuscript) names of books and periodicals, bulletins, etc. (*See examples below.*)

Uniform sequence of data shoud be observed, as follows: *For a book*—Author's name (last name first); title of article or chapter; name of book; place of publication; publisher; year date. *For a periodical*—Author's name (last name first); title of article or chapter; name of periodical; volume number; inclusive page numbers; year date.

Examples of References Appearing at End of Article:

1. American Association for Health, Physical Education, and Recreation. Suggested platforms for health education. *Journal of the American Association for Health-Physical Education-Recreation* **18**:436 (Sept. 1947).

2. American Association of School Administrators. *Health in Schools.* Revised edition. Washington, D. C.: the Association, a department of the National Education Association. pp. 266-7.

3. Deaver, G. G., Exercise and Heart Disease. *Research Quarterly*, **26**: 24-34, 1939.
4. Ogden, Jean, and Jess Ogden. Small Communities in Action. New York: Harper & Brothers, 1946.
5. Potter, John Nicholas. *Physical Fitness of Junior High School Boys.* Unpublished Master's thesis, University of California, Berkeley, 1942.

Tabular Matter

Each table should have a descriptive heading and should be specifically referred to in the text by number, *e.g.*, "Table 1[illegible]" etc., never as "the above table" or "the following table." Number tables from 1 up for the entire manuscript, using Arabic numerals. Do not duplicate data by giving it in *both* tables and graphs.

Tables should be double-spaced typewritten, like the rest of the material in the manuscript. They should be typed on separate sheets, as the printer will set them on a different machine from the one used for the text matter. If a table continues on a second sheet, it is not necessary to repeat the boxheads, since the printer will repeat from the original boxheads, when necessary.

The word "TABLE" should be written in capital letters, as: "TABLE 1"; the table title should be written in lower case letters with initial capitals, and centered over the table. Tables should be ruled as desired, except that no rules will appear at the extreme right and left edges of the table. No double rules are to be used, unless necessary for clarity.

Well-known statistical formulas should be omitted. Extensive tabular material, raw data, and appendixes should not be printed; the author can mention in a footnote that he will supply such material in mimeographed form on request.

Illustrations

Illustrative material is of two types: pen and ink drawings, which are reproduced by the line engraving process; and photographs, wash drawings, stipple drawings (in short, anything containing shading), which are reproduced by the halftone process.

Line engravings are always treated as text figures and should be so designated. All drawings should be made with India ink, preferably on white bristol board plate, 1 ply or 2 ply, which is sufficiently transparent to permit tracing if back lighting (e.g., a window pane) is used. Avoid graph paper for the reproduction copy, as the printing interferes with proper inking and the paper permits no corrections. Sometimes it is desirable to ink in the principal guide lines so that the curves can be more easily read. Good examples of graphs can be seen in *The Research Quarterly* for October 1953, pages 332 and 366.

Lettering should be plain and large enough to reproduce well when the drawing is reduced to the dimensions of the printed page (4⅛ x 7 inches). Most figures can be advantageously drawn for a linear reduction of one-half or one-third. Be sure to draw the lines heavy enough so that they will not be overly thin after reduction. Explanatory lettering should be included within the chart. Typewritten lettering does not reproduce well; it is much better to use a LeRoy or similar lettering device.

Care should be taken not to waste space, as this means greater reduction and a less satisfactory illustration. Often it is possible to combine several curves in one figure and enable the reader to make comparisons.

Halftones are treated as figures and should be so designated. Frequently, several halftones can be grouped to form an attractive full page, in which case they should be numbered consecutively, in Roman numerals. Photographs should be in the form of clear black-and-white prints on glossy paper. Care should be taken to see that they cannot be bent or folded in handling *and paper clips should not be used.* All imperfections are reproduced.

Figures should each be numbered consecutively from I up for the entire manuscript. Use Roman numerals to number figures, and Arabic numerals to number tables. The legends for the illustrations should be typed upon a separate sheet placed at the end of the manuscript. Care should be taken to indicate plainly in the text the exact location of all illustrations and tables.

The Association will assume complete engraving expense.

Special Points of Style

USE OF NUMBERS

Use Arabic figures for all definite weights, measurements, percentages, and degrees of temperature (for example: 2 kgm., 1 inch, 20.5 cc., 300° C.). Spell out all indefinite and approximate periods of time (for example: over one hundred years ago, about two-and-one-half hours). For numerals used in a general sense, spell out numbers through ten and use Arabic figures for 11 and over (seven times, five years old, 11 students).

ABBREVIATIONS AND SYMBOLS

Standard abbreviations should be used whenever the weights and measurements are used with figures, i.e., 10 kg., 6.25 cc., etc. The forms to be used (for both singular and plural) are: ft., ft.-lb., ft/sec, in., yd., min., hr., sq. ft., sq. in., rpm. *Gram* should be spelled out in all cases to avoid possible confusion with *grain;* also spell out *mile.* All obscure and ambiguous abbreviations should be avoided. Symbols used should follow the notation listed in *Research Methods* (AAHPER), pp. 518-20 and 522-25. Most common are:

M = mean	r = Pearson correlation
Mdn = median	r_{bis} = biserial correlation
N = number of individuals	r_{11} = reliability coefficient
n = number of measurements	χ^2 = chi square
σ = standard deviation	F = variance ratio
σ_M = standard error of mean	t = Student (Fisher) t ratio

Per cent should be two words. Use per cent sign (%) in tables or when it appears in parentheses in text.

Proofreading

The author will receive his original manuscript and any engraver's proofs with the galley proofs of his article for correction. A reprint order blank will be enclosed for the author's convenience.

Corrected proofs and original manuscripts are to be returned within 48 hours by first-class mail to the Editor, AAHPER, 1201 Sixteenth Street, Northwest, Washington 6, D. C.

RESEARCH HELPS

RESEARCH METHODS
(Revised Edition)
535 pp. **$5.00**

MEASUREMENT AND EVALUATION MATERIALS
138 pp. **$2.50**

Prepared by the National Research Council
of the Research Section of the AAHPER

MASTERS THESES
1930-46
Over 3,500 titles
cross-indexed under subjects
and areas
292 pp. **$3.00**

DOCTORATE THESES
1930-46
List of 420 theses
reported by graduate
departments
40 pp. **50c**

by Thomas K. Cureton
University of Illinois

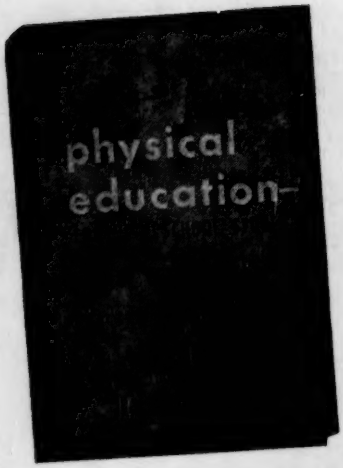
RESEARCH UNDERWAY and RESEARCH NEEDED
in
Health Education, Physical Education, and Recreation

Survey directed by Research Council of AAHPER
56 pp., mimeo. $1.00

Order these professional publications from

AMERICAN ASSOCIATION FOR HEALTH,
PHYSICAL EDUCATION, AND RECREATION

1201 Sixteenth St., N.W. **Washington 6, D. C.**



PHYSICAL EDUCATION FOR HIGH SCHOOL STUDENTS

"The publication of this book marks one of our greatest professional achievements."
—Eleanor Metheny

A book of sports, games, athletics, and recreation activities for teen-agers

Clothbound cover—$3.00

416 pages 6" x 9"

Cover in four colors

NEA discount on quantity orders: 2-9 copies, 10%
10 copies or over, 20%

AAHPER Publications-Sales
1201 - 16th St., N.W.
Washington 6, D. C.

CONTENTS

PART I—GETTING STARTED
What Physical Education Can Do for you—Keeping Fit.

PART II—ACTIVITIES
Sports
Angling — Archery — Badminton — Baseball and Softball — Basketball — Bowling — Field Hockey — Golf — Gymnastics and Tumbling — Handball — Lacrosse — Riflery — Soccer — Speedball — Swimming — Tennis — Touch Football — Track and Field — Volleyball — Wrestling.
Dance
Folk Dance — Social Dance — Square Dance — Modern Dance.
Games and Parties
Recreational Games — Co-recreational Parties.

PART III—ATHLETICS
Intramurals — Interscholastics — Keeping in Top Condition.

PART IV—CAREER INFORMATION
Physical Education as a Career — Recreation as a Career.